

UNIVERSIDAD COMPLUTENSE DE MADRID

FACULTAD DE CIENCIAS MATEMÁTICAS



TESIS DOCTORAL

Aproximación Bayesiana aplicada al Reparto Modal en Modelos de Transporte

MEMORIA PARA OPTAR AL GRADO DE DOCTOR

PRESENTADA POR

Óscar de Gregorio Vicente

DIRECTORES

Miguel Ángel Gómez Villegas
Beatriz González Pérez

UNIVERSIDAD COMPLUTENSE DE MADRID
FACULTAD DE CIENCIAS MATEMÁTICAS



TESIS DOCTORAL

Aproximación Bayesiana aplicada al Reparto Modal en
Modelos de Transporte

MEMORIA PARA OPTAR AL GRADO DE DOCTOR
PRESENTADA POR

Óscar De Gregorio Vicente

DIRECTORES

Miguel Ángel Gómez
Villegas (UCM)

Beatriz González Pérez
(UCM)

Año 2020

Programa de Doctorado en Ingeniería Matemática,
Estadística e Investigación Operativa por la
Universidad Complutense de Madrid y la
Universidad Politécnica de Madrid



Aproximación Bayesiana aplicada al Reparto Modal en Modelos de Transporte

TESIS DOCTORAL

Óscar De Gregorio Vicente

DIRECTORES

Miguel Ángel Gómez
Villegas (UCM)

Beatriz González Pérez
(UCM)

Año 2020



U N I V E R S I D A D
COMPLUTENSE
M A D R I D

**DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD DE LA TESIS
PRESENTADA PARA OBTENER EL TÍTULO DE DOCTOR**

D. *Oscar De Gregorio Vicente*, estudiante en el Programa de Doctorado de *Ingeniería Matemática, Estadística e Investigación Operativa (IMEIO)*, de la Facultad de *Ciencias Matemáticas* de la Universidad Complutense de Madrid, como autor de la tesis presentada para la obtención del título de Doctor y titulada:

Aproximación Bayesiana aplicada al Reparto Modal en Modelos de Transporte

y dirigida por:

Miguel Ángel Gómez Villegas y Beatriz González Pérez

DECLARO QUE:

La tesis es una obra original que no infringe los derechos de propiedad intelectual ni los derechos de propiedad industrial u otros, de acuerdo con el ordenamiento jurídico vigente, en particular, la Ley de Propiedad Intelectual (R.D. legislativo 1/1996, de 12 de abril, por el que se aprueba el texto refundido de la Ley de Propiedad Intelectual, modificado por la Ley 2/2019, de 1 de marzo, regularizando, aclarando y armonizando las disposiciones legales vigentes sobre la materia), en particular, las disposiciones referidas al derecho de cita.

Del mismo modo, asumo frente a la Universidad cualquier responsabilidad que pudiera derivarse de la autoría o falta de originalidad del contenido de la tesis presentada de conformidad con el ordenamiento jurídico vigente.

En Madrid, a 25 de noviembre de 2020

Fdo.: Óscar De Gregorio Vicente

Esta DECLARACIÓN DE AUTORÍA Y ORIGINALIDAD debe ser insertada en la primera página de la tesis presentada para la obtención del título de Doctor.

A Marco

-¿Soñaré?

-Desde luego que soñarás.

Todas las criaturas inteligentes sueñan, pero nadie sabe por qué.

(Arthur C. Clarke, 2010, Odisea Dos)



UNIVERSIDAD COMPLUTENSE DE MADRID
Y UNIVERSIDAD POLITÉCNICA DE MADRID

APROXIMACIÓN BAYESIANA APLICADA AL REPARTO MODAL EN MODELOS DE TRANSPORTE

TESIS PRESENTADA POR ÓSCAR DE GREGORIO VICENTE
PARA OBTENER EL GRADO DE DOCTOR
DENTRO DEL PROGRAMA DE DOCTORADO EN
INGENIERÍA MATEMÁTICA, ESTADÍSTICA E INVESTIGACIÓN OPERATIVA

2020

Directores:

Miguel Ángel Gómez Villegas y Beatriz González Pérez

Agradecimientos

A mi mujer Teresa y a mis hijos Guillermo y Elisa, por su apoyo, ánimo, generosidad infinita y comprensión incondicional. Ha sido una larga aventura y ellos siempre han estado a mi lado.

A mis padres, abuelos y resto de mi familia, por darme las oportunidades y los valores necesarios y adecuados durante toda mi vida.

A mis tutores, el Dr. M. Ángel Gómez Villegas y la Dra. Beatriz González Pérez, por sus aportaciones y por su dedicación en la elaboración de este trabajo.

Al Grupo de Investigación en Métodos Bayesianos del Departamento de Estadística e Investigación Operativa de la Universidad Complutense de Madrid, por sus ideas en cuanto a las diferentes aplicaciones y líneas de investigación con Redes Bayesianas.

Al Grupo Integrado de Ingeniería -GII- de la Universidade da Coruña, en especial a Diego y Rosa, con quienes he compartido horas de trabajo, viajes y descubrí el salar de Uyuni.

Al resto de compañeros y amigos que de alguna forma han estado presentes durante estos años.

Esta investigación ha sido parcialmente subvencionada por el Ministerio de Economía y Competitividad, España, grant TRA2015-65283-R y la Universidad Complutense de Madrid–Banco Santander, España, grant PR26/16-20261.

La fuente de la figura de la portada es el trabajo de Salinas Roberto (2013). Recuperado del enlace <https://slideplayer.es/slide/2763492/>.

Resumen

El trabajo presentado en esta Memoria está orientado a demostrar que la utilización de las Redes Bayesianas (RB), aplicadas al área de la planificación de los transportes representa una mejora innovadora comparándola con otras técnicas habitualmente empleadas hasta la fecha como son los modelos Logit y las Redes Neuronales (RN).

En concreto, su utilización en la etapa del reparto modal de modelos de demanda de cuatro etapas (generación-atracción, distribución, reparto o elección modal y asignación), que es aquella correspondiente en la que se enfrenta la unidad de decisión dentro de un conjunto discreto de alternativas, tanto para pasajeros como mercancías.

Las preguntas fundamentales que se responden a lo largo de esta memoria son: ¿cuáles son las contribuciones originales de este trabajo?, ¿cuál es su naturaleza?, ¿cuál es su alcance? y ¿cuáles son sus limitaciones?.

Los resultados obtenidos en este trabajo no permiten una generalización, de momento, para modelos de transporte de mercancías, sino que son específicos para este caso. Teniendo en cuenta, además, que la confidencialidad del uso de los mismos condiciona el detalle de las alternativas propuestas y por tanto, de los resultados mostrados. Al respecto, mencionar que en el capítulo 4.4 Metodología, se enumeran otros trabajos similares recientes sobre modelos de transporte como los de Sun et al. (2006), Correa et al. (2009), Tang et al. (2012), Daziano et al. (2013), Yannis Tyrinopoulos (2013) y Tai Yu Ma (2015), con resultados igualmente prometedores. Sin embargo, todos ellos se focalizan en ámbitos de estudio urbanos o metropolitanos y de vehículo privado, por lo que difieren en cuanto a su uso hacia un modelo de transporte de mercancías, como el indicado en el caso de estudio que se presenta, el

cual además es un caso real.

Ésta es la principal contribución original de este trabajo, el uso de una RB en la etapa de reparto modal de un modelo clásico de cuatro etapas, aplicado a un modelo multimodal de mercancías de ámbito mundial.

Los objetivos generales que se han perseguido con este trabajo son: a.1) mejorar el proceso de decisión de viajeros y operadores de carga, en los procesos de participación pública y declaración de preferencias, a partir del análisis detallado de la información de partida. La aplicación de las RB para la obtención de resultados de cada alternativa, en comparación con las otras técnicas, se demuestra en este caso que ofrece mejores ajustes en las variables principales de tiempo y coste por alternativa; a.2) optimizar los modelos de transporte utilizados en planificación; a.3) comprobar que la aplicación de las aproximaciones bayesianas en la etapa de reparto modal supone obtener mejores resultados; a.4) promover las redes bayesianas en entornos complejos que ya son una realidad como son los Sistemas Inteligentes de Transporte (ITS), los Sistemas Cooperativos (V2V, V2I, V2I2V, I2I) de transporte e infraestructuras, y los vehículos autónomos.

Los objetivos específicos, por otro lado, se resumen en: b.1) conseguir mejores resultados con menor cantidad de información; b.2) disminuir los tiempos de procesamiento para el reparto modal; b.3) mejorar el ajuste coste-beneficio derivado de las mejoras aportadas por las RB en los resultados, cara a optimizar los recursos económicos existentes y conseguir financiación para los proyectos de infraestructura y transporte. En este caso hay que mencionar que el detalle de los costes de cada alternativa de transporte presentada está sujeto a confidencialidad. Sin embargo, en enlace http://www.iirsa.org/proyectos/detalle_proyecto.aspx?h=1351 se muestra la ficha descriptiva del proyecto correspondiente al COSIPLAN (Consejo Suramericano de Infraestructura y Planeamiento) de UNASUR (Unión de Naciones Suramericanas). Comprobándose que es uno de los proyectos más importantes del IIRSA (Iniciativa para la Integración de la Infraestructura Regional Suramericana) con una inversión total de unos 7.000 millones de dólares; b.4) mejorar los resultados respecto a los algoritmos utilizados habitualmente.

El caso de estudio presentado está basado en el proyecto de un corredor ferroviario en Sudamérica, en el que partiendo del modelo de transporte asociado se obtienen

resultados comparativos de los errores correspondientes a la elección de las posibles alternativas de transporte de mercancías, aplicando los tres métodos indicados: un Multi Nomial Logit (MNL), una Red Neuronal y una Red Bayesiana.

Los resultados obtenidos permiten concluir que la aproximación bayesiana consigue los mejores ajustes de las tres técnicas, minimizando los errores en la elección del modelo de reparto modal, por lo que se demuestra la hipótesis de partida.

En cuanto a las futuras líneas de investigación: mejora de los modelos macroscópicos de 4 etapas en las otras etapas en fase de estudio, fundamentalmente la etapa 1-Generación-Atracción y la etapa 2-Distribución; mejora de los algoritmos de seguimiento vehicular en los modelos microscópicos e híbridos de simulación de tráfico, partiendo de datos obtenidos en tiempo real de los centros de control; optimización en las técnicas matemáticas empleadas en los modelos de eventos discretos, en relación al tratamiento de las longitudes máxima de cola y los servidores, mejorando los tiempos en los servicios; Sistemas Expertos aplicados al COVID19; mejora de las técnicas aplicadas en Digital Humanities en el procesamiento de la información obtenida de imágenes de alta resolución; Sistemas Inteligentes de Transporte (ITS) y Sistemas Cooperativos vehículo-infraestructura; vehículos de conducción autónoma, RPAS (drones) y sistemas robóticos; redes sociales, redes de telecomunicación, redes de abastecimiento y distribución (electricidad, agua, gas. . .) en cuanto al ajuste de los modelos de consumo y de demanda; modelos genéticos, modelos tumorales, innovación médica.

Abstract

The work showed in this Report is aimed at demonstrating that the use of Bayesian Networks (BN), applied to the area of transport planning, represents an innovative improvement compared to other techniques commonly used to date such as Logit models and Neural Networks (NN).

Specifically, its use in the stage of modal split of four-stage demand models (trip generation, distribution, modal split and assignment), which is the one in which the decision unit is faced within a discreet set of alternatives, both for passengers and freight.

The key questions that are answered throughout this Report are: what are the original contributions of this work?, what is its nature?, what is its scope? and what are its limitations?.

The results obtained in this work do not allow a generalization, for the moment, for freight transport models, but are specific for this case. Taking into account, in addition, that the confidentiality of their use conditions the detail of the proposed alternatives and, therefore, of the results shown. In this regard, it should be mentioned that in chapter 4.4 Methodology, other similar recent works on transport models such as those of Sun et al. (2006), Correa et al. (2009), Daziano et al. (2013), Yannis Tyrinopoulos (2013) and Tai Yu Ma (2015), with equally promising results. However, all of them are focused on urban or metropolitan areas of study and private vehicles, so they differ in terms of their use towards a freight transport model, such as that indicated in the case study presented, which it is also a real case.

This is the main original contribution of this work, the use of a BN in the modal split stage of a classic four-stage model, applied to a world-wide multimodal freight model.

The general objectives that have been pursued with this work are: a.1) improving the decision-making process of travelers and cargo operators, in the processes of public participation and declaration of preferences, based on the detailed analysis of the initial information. The application of BN to obtain results for each alternative, in comparison with the other techniques, is shown in this case to offer better adjustments in the main variables of time and cost per alternative; a.2) optimize transport models used in planning; a.3) checking that the application of Bayesian approximations in the modal split stage provides better results; a.4) promoting Bayesian networks in complex environments that are already a reality such as Intelligent Transportation Systems (ITS), Cooperative Systems (V2V, V2I, V2I2V, I2I) of transport and infrastructures, and autonomous vehicles.

The specific objectives, on the other hand, are summarized as follows: b.1) achieving better results with less information; b.2) decreasing processing times for modal split; b.3c) improving the cost-benefit adjustment derived from the improvements contributed by the BRs in the results, in order to optimize existing economic resources and obtain financing for infrastructure and transport projects. In this case, it should be mentioned that the detail of the costs of each transport alternative presented is subject to confidentiality. However, the following official link shows the descriptive file of the project corresponding to the COSIPLAN (South American Council for Infrastructure and Planning) of UNASUR (Union of South American Nations): http://www.iirsa.org/proyectos/detalle_proyecto.aspx?h=1351. Proving that it is one of the most important projects of the IIRSA (Initiative for the Integration of Regional Infrastructure in South America) with a total investment of about 7,000 million USD; b.4) improving the results regarding commonly used algorithms.

A case study based on the project of a railway corridor in South America is presented, in which, starting from the associated transport model, comparative results of the errors corresponding to the choice of possible merchandise transport alternatives are obtained, applying the three methods indicated: a Multi Nomial Logit (MNL), a Neural Network and a Bayesian Network.

The results obtained lead to the conclusion that the Bayesian approach provides the best adjustments of the three techniques, minimizing the errors obtained in the

choice of the modal split model, thus proving the starting hypothesis.

Regarding future lines of research: improvement of the 4-stage macroscopic models in the other stages under study, fundamentally stage 1-Generation-Attraction and stage 2-Distribution; improvement of vehicle tracking algorithms in microscopic and hybrid traffic simulation models, based on data obtained in real time from control centers; optimization of the mathematical techniques used in the discrete event models, in relation to the treatment of maximum queue lengths and servers, improving service times; Expert Systems applied to COVID19; improvement of the techniques applied in Digital Humanities in the processing of information obtained from high resolution images; Intelligent Transportation Systems (ITS) and Vehicle-Infrastructure Cooperative Systems; autonomous driving vehicles, RPAS (drones) and robotic systems; social networks, telecommunication networks, supply and distribution networks (electricity, water, gas. . .) in terms of adjusting consumption and demand models; genetic models, tumor models, medical innovation.

Prólogo

El presente trabajo nace a partir de la idea de que los avances en los últimos diez años en el sector de la movilidad y el transporte, tanto en entornos urbanos como en entornos regionales o mundiales, requieren innovar y aplicar nuevas soluciones a los modelos de transporte utilizados hasta la fecha, tanto para pasajeros como para mercancías.

Hay que pensar, por un lado, que los Sistemas Inteligentes de Transporte (ITS) y las herramientas de modelización están llevando un avance paralelo en soluciones innovadoras que, en algún momento, les hará converger de forma integral.

Por ahora, esa integración se está realizando a nivel modular con algunas aplicaciones de ITS, que se están viendo reflejadas en la calidad de los modelos de transporte urbanos, ya que, muchos de los datos de entrada de estos mismos llegan en tiempo real a través de sensores o cámaras.

Este avance, está permitiendo poco a poco, disponer de herramientas de planificación de transporte que son de dos tipos: interactivas, con los objetos con los que se relacionan; iterativas, a nivel de las soluciones alcanzadas.

Por otro lado, la Planificación de los Transportes de una determinada zona necesita herramientas que permitan representar o modelizar, de forma analítica, la situación actual y futura de su red de infraestructuras.

Una de estas herramientas son los denominados Modelos de Transporte que son representaciones virtuales de una red física, formada por arcos y nodos que poseen unas determinadas características, en la que participan diferentes modos de transporte.

Los objetivos principales de estos modelos son: investigar el comportamiento actual de pasajeros y mercancías; predecir ese comportamiento a futuro, bajo dife-

rentes hipótesis que configuran diferentes escenarios; y aplicar los resultados como fundamento para la identificación y evaluación de problemas y la planificación de actuaciones.

De esta manera, no sólo se pueden estimar demandas futuras para cada uno de los escenarios considerados, si no que derivados de estos resultados se pueden calcular las inversiones necesarias en transporte (material rodante o unidades), infraestructuras (red de carreteras, puertos, aeropuertos, ferrocarriles, hidrovías, corredores marítimos, etc.), servicios, desarrollos urbanísticos, políticas sociales, políticas tarifarias, etc. . . Así como los costes e ingresos, directos e indirectos, lo que permite obtener las rentabilidades o beneficios globales de las actuaciones consideradas, marcando una matriz de jerarquías a nivel estructural, que permita ayudar en la toma de decisiones de alto nivel.

Parte de los resultados incluidos en esta Memoria ya han sido publicados en un artículo sujeto a J.C.R. de De Gregorio-Vicente, Ó (septiembre 2017).

Existen dos tipos principales de Modelos de Transporte: los basados en viajes, siendo la unidad un viaje entre un origen y un destino; los basados en actividades, donde se analiza la cadena de viajes en un día completo, derivada de realizar una serie de actividades.

A los primeros se les denomina modelos clásicos de cuatro etapas (four-step algorithm), mientras que a los segundos se les denomina (tour-based model). Si bien las máquinas de procesamiento actuales son capaces de trabajar con ambos, a nivel algorítmico, el principal problema reside en la calibración de los segundos, ya que requiere de mucha información para cada una de las etapas consideradas, lo cual supone un diferencial económico a considerar (por tiempo y personal) respecto de los primeros. Por tanto, el presente trabajo se centra en la utilización de los modelos clásicos, cuyas cuatro etapas, en cuanto a los viajes modelizados en la red, son: Generación-Atracción (G/A); Distribución (D); Reparto Modal (RM); y Asignación a la red (A).

Los niveles de estudio o detalle de estos modelos, a su vez, se pueden definir en: macroscópico, mesoscópico y microscópico. Pudiendo darse la situación de que determinados softwares permiten la visualización simultánea de un nivel macroscópico y un nivel microscópico, en lo que se denominan modelos híbridos de representación.

Dicho esto, actualmente nos encontramos con tres factores fundamentales que nos han de replantear la forma de abordar o construir los modelos de transporte: la unidad de decisión (viajero o cargador) no es Ideal y objetiva, sino que es Real y subjetiva, y parte de información previa para la toma de decisiones; los Sistemas Cooperativos, basados en las comunicaciones e intercambio de información: entre vehículos (V2V), entre vehículo e infraestructura (V2I), entre vehículo-infraestructura-vehículo (V2I2V) y dentro de la misma infraestructura (I2I), son ya un hecho; los vehículos de conducción autónoma se impondrán en el corto-medio plazo, lo que afectará de lleno a la planificación de los transportes e infraestructuras.

Aunque recoger todas las citas es prácticamente imposible, debemos citar a Ricardo A. Daziano, Luis Miranda-Moreno & Shahram Heydari (2013), donde están contenidas las RB asociadas a la distribución multinomial; y también citar a Tai-Yu Ma (2015) donde pueden verse distribuciones más generales aplicadas a este tipo de redes

La distribución de la presente Memoria es la siguiente: en la sección 1 se explican los conceptos fundamentales de la Teoría de Grafos, Probabilidad e Independencia; en la sección 2 se presentan las Redes Bayesianas, en cuanto a sus antecedentes, definición, propiedades, tipología, aplicaciones y el detalle de las Redes Bayesianas Gaussianas; la sección 3 explica los puntos fundamentales de los Modelos de Transporte, los fundamentos básicos, los modelos de cuatro etapas, y las consideraciones en cuanto a los modelos de elección o reparto modal; la sección 4 describe un caso de estudio acerca de un corredor ferroviario, principalmente, de mercancías, incluyendo la construcción y calibración de un MultiNomial Logit (MNL), una Red Neuronal (RN) y una Red Bayesiana (RB) ajustada al caso concreto; la sección 5 incluye las principales conclusiones obtenidas del trabajo; la sección 6 muestra las futuras líneas de investigación; la sección 7 la bibliografía principal consultada, si bien la documentación global es más extensa; la sección 8 incluye los Anexos principales, formados por el conjunto de datos de partida (Data Set), así como las macros utilizadas en R para la construcción de las Redes empleadas.

Índice general

Agradecimientos	2
Resumen	3
Abstract	6
Prólogo	9
1. Conceptos fundamentales	20
1.1. Fundamentos de la Teoría de Grafos	20
1.1.1. Antecedentes históricos	20
1.1.2. Conceptos básicos de Grafos	22
1.1.3. Tipos de Grafos	24
1.1.4. Grafos dirigidos	26
1.1.5. Otras estructuras gráficas	28
1.2. Probabilidad e independencia	30
1.2.1. Probabilidad condicionada	31
1.2.2. Independencia	31
1.2.3. Independencia condicionada	32
1.2.4. Independencia condicionada en variables aleatorias	33
1.2.5. Propiedades de la independencia condicionada en variables aleatorias	34
1.2.6. D-separación y evidencia	34

2. Redes Bayesianas (RB)	37
2.1. Antecedentes	37
2.2. Definición y propiedades	38
2.3. Tipología y aplicaciones	42
2.3.1. Redes bayesianas discretas (RBD)	42
2.3.2. Redes bayesianas gaussianas (RBG)	43
2.3.3. Redes bayesianas mixtas (RBM)	47
2.3.4. Aplicaciones	48
2.4. Redes Bayesianas Gaussianas (RBG)	49
2.4.1. Construcción de una RBG	50
2.4.2. Independencia e Independencia condicionada	52
2.4.3. Propagación de la evidencia en RBG	53
2.4.4. Normalidad y No Normalidad	54
2.4.5. Test de Doornick-Hansen	55
2.4.6. Distribución Nonparanormal	57
3. Modelos de transporte	60
3.1. Introducción	60
3.2. Fundamentos previos	61
3.2.1. ¿Qué son y para qué se usan?	61
3.2.2. Objetivos principales	61
3.2.3. Tipología y niveles de detalle	63
3.3. Modelo clásico de cuatro etapas	64
3.3.1. Introducción	64
3.3.2. Especificaciones del Modelo clásico de cuatro etapas	66
3.3.3. Etapa 1: Generación-Atracción de viajes.	70
3.3.4. Etapa 2: Distribución de viajes	74
3.3.5. Etapa 3: Reparto o elección modal	82
3.3.6. Etapa 4: Asignación a la red	88
3.3.7. Métodos de asignación	90
3.4. Consideraciones de los modelos de elección modal	92
3.4.1. Multinomial Logit (MNL)	96
3.4.2. Redes Neuronales (RN)	98

4. Caso de estudio: Corredor Ferroviario Bioceánico Central	103
4.1. Introducción	103
4.2. Objetivos	103
4.2.1. Objetivos generales	103
4.2.2. Objetivos específicos	104
4.3. Marco de actuación	104
4.4. Metodología	107
4.5. Base de datos	111
4.6. Alternativas de elección	112
4.7. Ficheros principales de datos y estructura	113
4.7.1. Archivo Excel Dataset Eleccion Modal	118
4.7.2. Hoja Excel DATASET Elección Modal	118
4.7.3. Archivo Excel RDataset	119
4.7.4. Fichero de texto elecmod	120
4.8. Propósitos del modelo de elección modal	120
4.9. Calibración del MNL	121
4.10. Calibración de la RN	128
4.11. Calibración de la RB	133
4.12. Ajuste en RStudio de RN y RB	138
4.12.1. Ajuste de redes neuronales	138
4.12.2. Ajuste de redes bayesianas	141
4.13. Resultados	142
5. Conclusiones	146
6. Futuras líneas de investigación	159
7. Bibliografía	162
A. Muestra de la Base de Datos	170
B. DATASET ELECCION MODAL	171
C. RDATASET: DATASET	173

D. RDATASET: DATANORM	175
E. MACRO: R_Red Neuronal.R	177
F. MACRO: Trainnet.R	178
G. MACRO: R_Red Bayesiana.R	179

Índice de figuras

1.1. Localización de Königsberg	21
1.2. Puentes de Königsberg	21
1.3. Grafo de los Puentes de Königsberg	22
1.4. Representación gráfica de una red social. Fuente: INESEM	23
1.5. Representación gráfica de los ISP y nodos de Internet. Fuente: jurvetson @ Flickr. Licencia: CC BY 2.0	24
1.6. Arista dirigida	25
1.7. Arista no dirigida	25
1.8. Grafo dirigido	25
1.9. Grafo no dirigido	26
1.10. Grafo mixto	26
1.11. Grafo o árbol simple (izquierda) y poliárbol (derecha)	27
1.12. Grafo cíclico dirigido (DCG) (izquierda) y Grafo acíclico dirigido (DAG) (derecha)	28
1.13. Grafo moral G^m del grafo no dirigido de la Figura 1.9	29
2.1. De arriba a abajo: conexión serial, conexión convergente y conexión divergente	41
2.2. Red con propiedad de Markov	42
2.3. Red Bayesiana Discreta	43
2.4. Red Bayesiana Gaussiana (desplazamientos vehiculares laterales con aceleración	46

3.1. Modelo macroscópico de Estados Unidos de América (Fuente: Caliper Corp.)	61
3.2. Modelo macroscópico de Madrid (Fuente: Ayuntamiento de Madrid (Central de Movilidad))	62
3.3. Desagregación de las diferentes etapas de viaje a lo largo de un día (Fuente: “Travel Demand Modelling” (Moshe Ben Akiva) – Massachusetts Institute of Technology (MIT))	63
3.4. Modelo clásico de las cuatro etapas (Ortúzar and Willumsen 2011)	65
3.5. Red multimodal	69
3.6. Matriz de viajes. Fuente: Modelling Transport Manual (Ortúzar, J.D., y Willumsen, L. G. (2011))	75
3.7. Modelo del perceptrón simple con una entrada fija (arriba). Red neuronal multicapa. Fuente: Ramírez, J.A. (abajo)	100
4.1. Desconexión ferroviaria red oriental y red andina	105
4.2. Propuesta de conexión con CFBC	105
4.3. Modelo de transportes (mundial)	107
4.4. Esquema relacional del modelo de transportes	110
4.5. Hoja Excel DATASET Eleccion Modal	114
4.6. Estructura DATA1	115
4.7. DATANORM	116
4.8. Fichero de texto elecmod	117
4.9. Ejemplo de comparativa situación base (arriba) y situación futura (abajo)	127
4.10. Ejemplo de muestra de datos de exportaciones e importaciones	129
4.11. Perceptron multicapa entrenada por el conjunto de datos	130
4.12. Grafo de la Red Bayesiana	136
4.13. Interfaz RStudio	139
4.14. Instalación librería <i>neuralnet</i>	139
4.15. Error cuadrático medio por alternativa de transporte	144
4.16. Error máximo absoluto por alternativa de transporte	145
B.1. DATASET ELECCION MODAL	172

C.1. RDATASET: DATASET 174

D.1. RDATASET: DATANORM 176

Índice de cuadros

3.1. Variables de las rutas	69
4.1. Alternativas disponibles en el modelo (C=carretera, MM=multimodal carretera/tren). * Esta alternativa no está disponible en la actualidad, sino que será fruto de la construcción del CFBC.	113
4.2. Parámetros función de utilidad / p-valor del test de significación . . .	124
4.3. Resultados ajuste MNL	126
4.4. Variables de entrada de la RN	129
4.5. Resultados ajuste RN	131
4.6. Variables para la Red Bayesiana	134
4.7. Resultados de la calibración de la RB	135
4.8. Resultados ajuste RB	135
4.9. Resultados comparativos ajuste MNL, RN y RB	143

Capítulo 1

Conceptos fundamentales

Este Capítulo recoge conceptos fundamentales que permiten identificar la Teoría de Grafos y Probabilidad, de la que derivan los modelos gráficos probabilísticos (o redes probabilísticas), a los que pertenecen las Redes Bayesianas (RB), las cuales se construyen a partir de aspectos cualitativos y cuantitativos. Muchos de esos conceptos matemáticos fundamentales básicos y de las redes bayesianas mostrados en esta memoria han sido recogidos previamente en otras investigaciones anteriores, entre las que destacan como referentes cercanos las de Gómez-Villegas, M.A., Maín, P. y Susi, R. (2006) y Gómez-Villegas, M.A., Maín, P. y Viviani, P. (2014), por lo que deben ser ya conocidos para el lector.

La Teoría de Grafos, permite representar y explicar el aspecto cualitativo del modelo en cuanto a la estructura de dependencia (o independencia) de un conjunto de variables aleatorias, y la Probabilidad, a su vez, el aspecto cuantitativo definiendo numéricamente estas relaciones.

1.1. Fundamentos de la Teoría de Grafos

1.1.1. Antecedentes históricos

La ciudad de Königsberg (hoy Kaliningrado) ha pasado a la historia por ser el lugar donde se planteó el famoso problema de sus siete puentes, y cuya solución dio lugar a la Teoría de Grafos.



Figura 1.1: Localización de Königsberg



Figura 1.2: Puentes de Königsberg

Para hacernos una idea, Königsberg fue la capital de Prusia Oriental desde la Baja Edad Media hasta 1945, cuando fue tomada por la actual Rusia y renombrada como Kaliningrado, capital del actual óblast de Kaliningrado. Fue fundada por la Orden Teutónica y tuvo gran relevancia estratégica como ciudad portuaria al situarse en la desembocadura del río Pregolia, que llega hasta la laguna del Vístula, comunicado a su vez con el mar Báltico por el estrecho de Baltisk.

En dicha ciudad, en el siglo XVIII había siete puentes (actualmente solo cinco) que cruzaban el río Pregolia y conectaban sus cuatro regiones (ver Figura 1.3), y los ciudadanos se preguntaban si con alguna ruta se podrían cruzar todos los puentes una y solo una vez para ahorrar tiempo en sus desplazamientos. Hay que hacerse una idea de que esto que para nosotros puede ser tan trivial, en aquella época no lo era

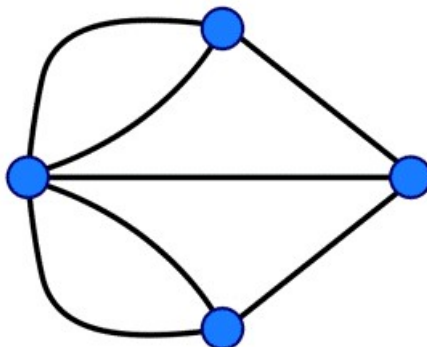


Figura 1.3: Grafo de los Puentes de Königsberg

tanto, habida cuenta de los precarios medios de transporte y vías de comunicación existentes entonces.

Leonard Euler (Basilea 1707, San Petesburgo 1783), halló la respuesta a esta pregunta y demostró que no era posible, modelizando la situación solo con aquello que era trascendente para el problema: las cuatro regiones (nodos) y los siete puentes (arcos) que las conectaban (ver Figura 1.3). La estrategia de resolución de este problema se considera la primera piedra de la Teoría de Grafos y de la Topología, al darse cuenta de que para que en un grafo de este tipo pueda recorrer cada arista, una y solo una vez, era necesario que a cada vértice llegara un número par de aristas, salvo a lo sumo dos vértices (inicial y final), y en este caso todos los vértices tendrían un grado impar, por lo que sería imposible cruzar todos los puentes una y solo una vez.

1.1.2. Conceptos básicos de Grafos

Un grafo representa gráficamente una colección de objetos $V = \{V_1, \dots, V_n\}$ llamados vértices (o nodos) relacionados entre sí mediante una selección de segmentos (aristas) que unen pares de vértices, que pueden ser dirigidos o no dirigidos. Típicamente los vértices se representan mediante puntos y las aristas mediante líneas. Siendo E_{ij} la arista que une los elementos V_i y V_j pertenecientes a V , se define implícitamente un grafo, donde $V = \{V_1, \dots, V_n\}$ son los nodos y E el conjunto de aristas que lo forman.

La importancia de los grafos reside en que son capaces de representar gráficamente

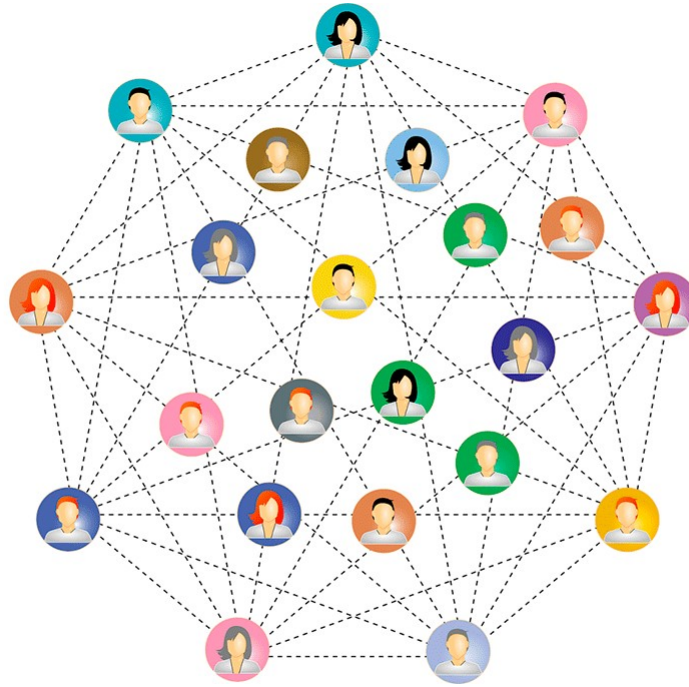


Figura 1.4: Representación gráfica de una red social. Fuente: INESEM

te, y de forma simplificada, problemas que se plantean en la vida real, como hemos visto en el apartado anterior, que pueden resolverse mediante diferentes técnicas y herramientas matemáticas de tal forma que es factible encontrar soluciones óptimas de forma más eficiente.

En la actualidad se pueden representar mediante grafos diferentes tipos de redes, entendiendo éstas como conexiones entre nodos mediante aristas, teniendo así redes de telecomunicación, redes de energía, redes de suministros, redes sociales, redes de transporte, etc. Por ese motivo, la teoría de grafos y su capacidad de representación, análisis y optimización de redes, mediante grafos, es una de las ramas de la matemática con mayor aplicabilidad y potencial a futuro, al utilizar combinatoria, álgebra, probabilidad, geometría, aritmética, etc.

En las Figuras 1.4 y 1.5 se muestran sendos grafos de una red social genérica, así como de los ISP (*Internet service provider*) y nodos de Internet. Ambas representaciones permiten modelizar de manera práctica las diferentes relaciones existentes en ambos casos, identificando los diferentes nodos o vértices y aristas en cada caso. Este tipo de representaciones es aplicable, como se ha indicado, a otro tipo de redes.

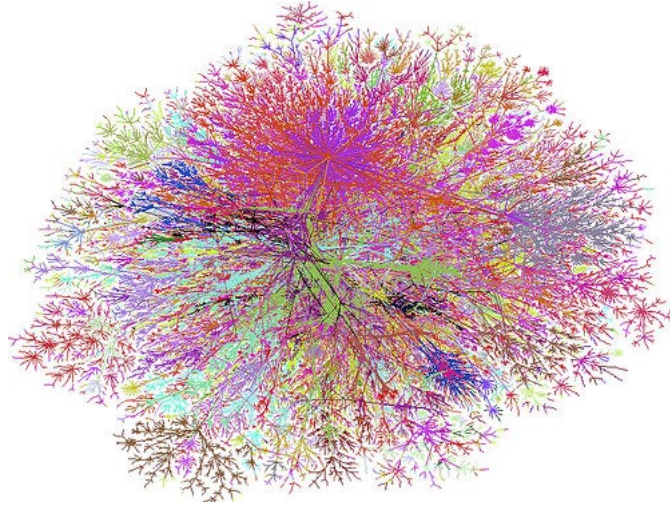


Figura 1.5: Representación gráfica de los ISP y nodos de Internet.
Fuente: jurvetson @ Flickr. Licencia: CC BY 2.0

Definición 1.1 Se define al **Grafo** $G=(V,E)$ como el par donde $V = \{V_1, V_2, \dots, V_n\}$ es un conjunto de elementos denominados vértices o nodos y $E \subseteq V \times V$ es un conjunto de aristas o arcos. Cada elemento de E , se representa por E_{ij} , indicando que corresponde a la arista que une V_i con V_j , para $i \neq j$.

1.1.3. Tipos de Grafos

Dependiendo de la relación y el orden que existe entre los nodos del grafo diferenciamos dos tipos de aristas: aristas dirigidas y aristas no dirigidas:

Aristas dirigidas, cuando $E_{ij} \in E$ pero $E_{ji} \notin E$ y se denotan como $V_i \rightarrow V_j$, tal que V_i se conecta con V_j y no viceversa. Es decir, representa una relación causa-efecto.

Aristas no dirigidas, se dan cuando $E_{ij} \in E$ y $E_{ji} \in E$ y se denotan como V_i-V_j , tal que ambos nodos V_i y V_j quedan conectados o relacionados de forma recíproca, quedando asociados entre sí.

En las Figuras 1.6 y 1.7 se representan ambos tipos de aristas.

A su vez, dependiendo de la cuantificación y tipo de aristas diferenciamos tres tipos de grafos:

Si un grafo tiene todas sus aristas del conjunto E dirigidas, entonces el grafo se define como **grafo dirigido** u orientado;



Figura 1.6: Arista dirigida



Figura 1.7: Arista no dirigida

Si un grafo tiene todas sus aristas del conjunto E no dirigidas, el grafo se define como **grafo no dirigido** o no orientado;

Si el grafo tiene aristas del conjunto E de forma que existen dirigidas y no dirigidas, el grafo se define como **grafo mixto** o de cadena.

En las Figuras 1.8, 1.9 y 1.10 se representan diferentes tipos de grafos.

Dado un grafo dirigido u orientado y una arista $\{v_1, v_2\}$, siendo v_1 el origen y v_2 es el fin de la arista, entonces las aristas así caracterizadas se llaman **flechas** del grafo.

Una arista de un grafo es un **lazo** si $\phi(A) = (v, v)$. En un grafo orientado una flecha es un lazo si no coinciden origen y fin.

Camino entre nodos es una sucesión de nodos conectados por una arista, tal que si se busca un camino entre los nodos V_i y V_j se tiene la sucesión de nodos $(V_{k_1}, \dots, V_{k_r})$ donde $V_i = V_{k_1}$ y $V_j = V_{k_r}$, de forma que existe una arista entre los nodos V_{k_l} y $V_{k_{l+1}} \forall l = 1, \dots, r$. El **camino es cerrado** cuando el nodo inicial del camino

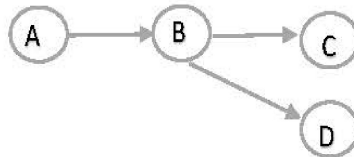


Figura 1.8: Grafo dirigido

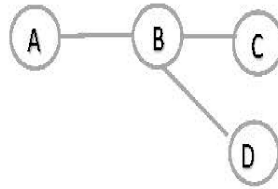


Figura 1.9: Grafo no dirigido

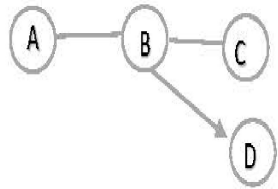


Figura 1.10: Grafo mixto

coincide con el nodo final del mismo, es decir, si $V_{k1} = V_{kr}$.

Se denomina **Ciclo** al camino cerrado en un grafo dirigido D , y se denomina **bucle** al camino cerrado en un grafo no dirigido.

Un **grafo G es conexo**, si para vértices v y w , tomados de dos en dos, existe una sucesión de lados l_1, \dots, l_k tal que v es adyacente a l_1 , w es adyacente a l_k , y cada dos lados l_i, l_{i+1} tienen un vértice adyacente en común. O dicho de otra forma, desde cualquier nodo inicial, existe al menos un camino que permite ir a todos los demás.

Las Redes Bayesianas (RB) son un caso específico de grafos dirigidos, y al ser éstas la base en la que se fundamenta la presente memoria, los análisis posteriores se centran únicamente en los grafos de esta tipología.

1.1.4. Grafos dirigidos

Las agrupaciones y relaciones entre los distintos nodos definen conjuntos de nodos concretos, como pasa también con los grafos no dirigidos, tal que, las relaciones familiares marcan algunas definiciones de igual carácter: familia, padre, hijo, ascendiente, descendiente, . . . De tal forma que:

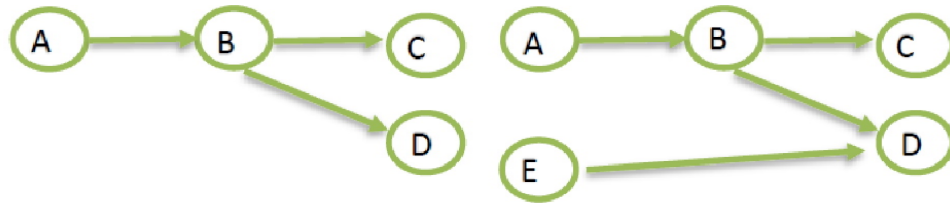


Figura 1.11: Grafo o árbol simple (izquierda) y poliárbol (derecha)

Si $V_i \rightarrow V_j$ se dice que V_i es *padre* de V_j , y se denota como $pa(V_j)$, y V_j es *hijo* de V_i .

El conjunto formado por la union de un nodo V_i y sus padres $pa(V_i)$ se le llama *familia del nodo* V_i , de forma que $fa(V_i) = V_i \cup pa(V_i)$.

El conjunto de todos los padres del nodo V_j se escribe como $pa(V_j)$.

El número de padres de un nodo define tipologías de grafos dirigidos: si cada nodo tiene como máximo un padre, el grafo dirigido se denomina **grafo o árbol simple**; en caso contrario **poliárbol**

Los ascendientes de un nodo V_i se denotan como $as(V_i)$, y son el conjunto de nodos que tienen un camino hasta V_i ; análogamente, son descendientes del nodo V_i , y se denotan $de(V_i)$ el conjunto de nodos a los que se puede ir desde V_i .

El conjunto de todos los nodos que no son ascendientes de V_i se denotan por $na(V_i) = V \setminus (as(V_i) \cup V_i)$. El conjunto de todos los nodos que no son descendientes de V_i se denotan $nd(V_i) = V \setminus (de(V_i) \cup V_i)$.

A partir de estos conceptos de nodos ascendientes y descendientes en un grafo dirigido $D = (V, E)$, surge identificar la ordenación entre dichos nodos. Asignando un número a cada nodo se obtiene una *numeración ancestral*, que es aquella tal que el número de cada nodo es menor que el correspondiente a sus hijos.

Tomando de ejemplo el poliárbol de la Figura 1.11, los nodos $\{B,C,D\}$ son descendientes del nodo A, siendo por su parte los nodos $\{A,B,E\}$ los ascendientes del nodo D . Por otro lado, $pa(D)=\{B,E\}$, y a su vez, $fa(D) = \{B,E,D\}$.

A continuación, se muestran algunas de las principales definiciones asociadas:

Definición 1.2 Un grafo dirigido $D = (V, E)$ es **acíclico (DAG, Directed Acyclic Graphs)** cuando no contiene ningún ciclo. Si contiene al menos un ciclo el grafo es un grafo **cíclico (DCG, Directed Cyclic Graphs)**.

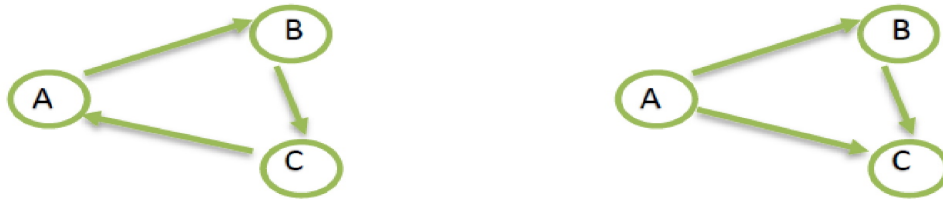


Figura 1.12: Grafo cíclico dirigido (DCG) (izquierda) y Grafo acíclico dirigido (DAG) (derecha)

Las RB se describen por grafos acíclicos dirigidos DAGs, que son grafos imprescindibles para poder especificar problemas en los que exista cierto grado de incertidumbre y en el que un conjunto de variables estén relacionadas. Siempre que existe un DAG existe un grafo no dirigido asociado que es el propio DAG en el que las aristas dirigidas se reemplazan por aristas no dirigidas.

Definición 1.3 Sea $G = (V, E)$ un grafo dirigido, entonces $G' = (V', E')$ es **subgrafo** de G si verifica que $V' \subseteq V$, $E' \subseteq E$ y G' es un grafo, tal que es un subconjunto de nodos, con sus enlaces de G .

Definición 1.4 Sea $G = (V, E)$ un grafo dirigido, se denomina **trayectoria** de G a una sucesión de nodos $V_1, V_2, \dots, V_{p-1}, V_p$, tales que $E_{i, i+1} \in E$, para cada $i = 1, 2, \dots, p - 1$.

La longitud de la trayectoria es el número de arcos recorridos para unir el nodo V_1 con el nodo V_p , siendo en ese caso de $p - 1$.

1.1.5. Otras estructuras gráficas

Además de estas definiciones existen estructuras gráficas más simples que se pueden obtener transformando un grafo, pero guardando ciertas propiedades, con el objetivo de simplificar el uso de éste en ciertos procedimientos, de cara a poder tratar los elementos del grafo localmente eliminando o minimizando posibles obstáculos computacionales. Estas estructuras son interesantes desde el punto de vista de las RB.

Definición 1.5 Sea $G = (V, E)$ un DAG, asociado a un grafo dirigido o un grafo

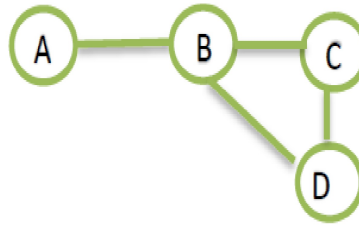


Figura 1.13: Grafo moral G^m del grafo no dirigido de la Figura 1.9

mixto, se denomina **grafo moral** G^m el obtenido uniendo con aristas no dirigidas a todos los nodos adyacentes no conectados con hijos comunes, y posteriormente, cambiando todos los enlaces dirigidos por conexiones no dirigidas.

También se puede obtener un grafo moral G^m a partir de un grafo no dirigido $G = (V, E)$, que puede venir asociado a un grafo dirigido o un grafo mixto.

Definición 1.6 Sea $G = (V, E)$ un DAG, asociado a un grafo dirigido o un grafo mixto, se denomina **esqueleto** de un grafo aquel que se obtiene sustituyendo la totalidad de los enlaces dirigidos por conexiones no dirigidas.

Además de estas definiciones, es interesantes tener en cuenta otras denominaciones como son la de grafo triangulado y la de grafo descomponible, obtenidos a partir de transformaciones de un grafo no dirigido, en el que: a la arista que articula dos nodos de un bucle existente que no es de dicho bucle se le nombra *cuerda del bucle*; y cuando a cada bucle de longitud mayor o igual a cuatro se le añade al menos una cuerda, resulta un *grafo triangulado*.

Definición 1.7 Sea $G = (V, E)$ un grafo no dirigido, se denomina **grafo triangulado** aquel en el que todos los bucles de longitud mayor o igual que cuatro contienen una cuerda, como mínimo.

Si se insertan cuerdas que dividen los bucles existentes en un grafo, entonces se puede triangular ese grafo. Esto es factible pero no es trivial, debido a que la estructura inicial del grafo debe permanecer invariante, por lo que la triangulación se realizará con el número mínimo de cuerdas, es decir, será una triangulación minimal. Y precisamente, una línea de investigación existente se basa en la construcción de algoritmos que garanticen que una triangulación sea minimal.

Definición 1.8 Sea $G = (V, E)$ un grafo no dirigido, tal que los subconjuntos dis-

juntos de nodos de V , no vacíos (A, B, C) , constituyen una **descomposición de G** si $V = A \cup B \cup C$, siendo C un subconjunto completo de V tal que cualquier camino de A a B , pasa por C .

Definición 1.9 Sea $G = (V, E)$ un grafo no dirigido, se le denomina **grafo descomponible** si: o es completo; o se obtiene una descomposición apropiada (A, B, C) siendo los subgrafos $G_{A \cup C}$ y $G_{B \cup C}$ grafos descomponibles.

De las anteriores definiciones de descomposición y triangulación de $G = (V, E)$, un grafo no dirigido, se obtiene el siguiente teorema obtenido de los resultados enunciados por Berge (1973) y Golumbic (1980):

Teorema 1.1 Sea $G = (V, E)$ un grafo no dirigido, son equivalentes las condiciones: G es un grafo triangulado; G es un grafo descomponible; cualquier separador mínimo existente entre dos vértices es completo.

1.2. Probabilidad e independencia

Conocidos los principales fundamentos de la Teoría de Grafos la cuestión es cómo poder relacionar un grafo con un conjunto de variables aleatorias, teniendo éstas su estructura de dependencia. La respuesta son las denominadas redes probabilísticas, o modelos gráficos probabilísticos, donde los nodos representan a las distintas variables y las aristas representan las relaciones de dependencia e independencia del conjunto de variables, tal que dichas relaciones de dependencia se dan con las distribuciones de probabilidad condicionadas.

A partir de las propiedades del grafo, en cuanto a sus nodos y aristas, se conoce la distribución de probabilidad conjunta de las variables, y a partir de dichas características es posible factorizar la distribución de probabilidad partiendo de otras funciones más sencillas, dadas para subconjuntos de variables. De esta forma es posible analizar las interdependencias de las variables. Las propiedades del grafo se denominan propiedades de Markov sobre grafos y consideran las diferentes tipologías de éstos

De esta forma surgen los modelos gráficos probabilísticos, uniendo la Teoría de

Grafos y la Teoría de la Probabilidad, lo que nos da una idea de las grandes capacidades matemáticas inherentes a ellos.

1.2.1. Probabilidad condicionada

Partamos de la definición de probabilidad condicionada e independencia de sucesos que se ven a continuación:

Definición 1.10 *Dados el espacio de probabilidad (Ω, Δ, P) y el suceso $A \in \Delta$ con $P(A) > 0$, se llama probabilidad condicionada del suceso $B \in \Delta$, dado el suceso A , al valor*

$$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

Teorema 1.2 *Sean (Ω, Δ, P) un espacio de probabilidad y $A \in \Delta$ tal que $P(A) > 0$. Entonces, (Ω, Δ, P_A) es un espacio de probabilidad, siendo*

$$P_A(B) = P(B|A), \quad B \in \Delta$$

1.2.2. Independencia

En cuanto a la independencia de sucesos, se han de tener en cuenta las siguientes definiciones:

Definición 1.11 *Se dice que dos sucesos $A, B \in \Delta$ son independientes si y sólo si*

$$P(A \cap B) = P(A)P(B).$$

Esta definición de independencia de los sucesos A y B también se cumple cuando $P(A) = 0$ y/o $P(B) = 0$. Además, si A y B son sucesos independientes, entonces A y B^c , A^c y B , así como A^c y B^c .

Considerando la condicionalidad de sucesos, otra definición de independencia de sucesos es la siguiente:

Definición 1.12 *El suceso A es independiente de B si*

$$P(A|B) = P(A)$$

Además, si A es independiente de B , también es cierto que B es independiente de A , es decir, se cumple la condición de simetría, por lo que:

$$P(B|A) = P(B)$$

De igual forma, si A y B no son independientes (son dependientes), entonces:

$$P(A|B) \neq P(A)$$

ó

$$P(B|A) \neq P(B)$$

Así, considerando las definiciones de independencia anteriores, y aplicando la regla de multiplicación, resulta que si A es independiente de B entonces,

$$P(A \cap B) = P(A|B) P(B) = P(A) P(B).$$

1.2.3. Independencia condicionada

Sean A , B y C tres sucesos, entonces se dice que A es condicionalmente independiente de B dado C , es decir, el suceso B no afecta al suceso A dado C si:

$$P(A|B \cap C) = P(A|C)$$

De forma similar a lo visto en el anterior apartado en relación a la propiedad de simetría de la independencia de sucesos, la independencia condicionada también es simétrica, tal que, si A es condicionalmente independiente de B dado C , entonces B es condicionalmente independiente de A dado C :

$$P(B|A \cap C) = P(B|C)$$

Igualmente, aplicando la regla de la cadena, se tiene que si A es condicionalmente independiente de B dado C , entonces:

$$P(A \cap B|C) = P(A|C) P(B|C)$$

En el caso de considerar $C = \Omega$, espacio muestral, podemos asumir que la independencia es un caso particular de la independencia condicionada.

1.2.4. Independencia condicionada en variables aleatorias

Las siguientes definiciones se basan en la Teoría de la Probabilidad y en los modelos gráficos probabilísticos en los que, como se indicó anteriormente, los nodos representan variables aleatorias y las aristas representan las relaciones de dependencia e independencia del conjunto de variables, tal que dichas relaciones de dependencia se dan con las distribuciones de probabilidad condicionadas.

Definición 1.13 Sean X e Y variables aleatorias; se dice que X e Y son variables aleatorias independientes, y se denota como $X \perp Y$, si y solo si

$$P(X, Y) = P(X) P(Y)$$

Esta igualdad es análoga a decir que su función de probabilidad conjunta es igual al producto de sus funciones de probabilidad marginales.

De la misma forma, basándonos en la probabilidad condicionada podemos expresar la independencia de las variables X e Y , tal que

$$P(X|Y) = P(X)$$

Teorema 1.3 Si $X \perp Y$, entonces $Y \perp X$.

Una vez vista la independencia de dos variables aleatorias dadas, X e Y , incluimos una tercera variable aleatoria, Z , para observar la independencia condicional entre ellas.

Lo anteriormente indicado nos lleva al siguiente teorema que resume el concepto de independencia de dos variables aleatorias X e Y .

Teorema 1.3 Si $X \perp Y$, entonces $Y \perp X$.

Una vez vista la independencia de dos variables aleatorias dadas, X e Y , incluimos una tercera variable aleatoria, Z , para observar la independencia condicional entre ellas.

Definición 1.14 Sean las variables aleatorias X , Y y Z , se dice que X e Y son condicionalmente independiente dado Z , y se denota como $X \perp Y | Z$, si y solo si

$$P(X|Y, Z) = P(X|Z)$$

Igualmente se cumple la independencia condicionada, $X \perp Y \mid Z$, si y solo si se cumple cualquiera de las siguientes condiciones:

1. $P(X, Y|Z) = P(X|Z) P(Y|Z)$, con $P(Z) > 0$
2. $P(X, Y, Z) = P(X|Z) P(Y|Z) P(Z)$, con $P(Z) > 0$
3. $P(X, Y, Z) = \frac{P(X,Z)P(Y,Z)}{P(Z)}$, con $P(Z) > 0$

1.2.5. Propiedades de la independencia condicionada en variables aleatorias

A continuación, se indican las principales propiedades de independencia condicionada. Estas propiedades son relevantes para conseguir nuevas relaciones entre variables, lo cual es importante dependiendo de la información de cada una, ya que esto está relacionado directamente con la estructura del modelo gráfico probabilístico en cuanto a sus nodos y aristas. Y eso es primordial no sólo para la construcción de los mismos, si no también para la interpretación derivada de las relaciones entre variables (o conjuntos de variables).

1. Conmutativa: si $X \perp Y \mid Z \Leftrightarrow Y \perp X \mid Z$.
2. Descomposición: si $X \perp (Y \cup W) \mid Z \Rightarrow X \perp Y \mid Z$ y $X \perp W \mid Z$.
3. Unión débil: si $X \perp (Y \cup W) \mid Z \Rightarrow X \perp Y \mid (W \cup Z)$ y $X \perp W \mid (Y \cup Z)$.
4. Contracción: si $X \perp Y \mid Z$ y $X \perp W \mid (Y \cup Z) \Rightarrow X \perp (Y \cup W) \mid Z$.
5. Intersección: si $X \perp Y \mid (Z \cup W)$ y $X \perp W \mid (Y \cup Z) \Rightarrow X \perp (Y \cup W) \mid Z$.

1.2.6. D-separación y evidencia

Tanto el concepto de **d-separación** como el de **evidencia** son fundamentales para entender las Redes Bayesianas, y por ese motivo se incluyen en este apartado las definiciones más representativas de ambos.

En primer lugar, se muestran dos definiciones fundamentales del concepto de *d*-separación gráfica de DAGs. La primera, de Pearl (1986) y la segunda, de Lauritzen, et al. (1990), siendo esta última la más utilizada.

Definición 1.15 (*d*-separación, Pearl) Sean X , Y y Z tres conjuntos disjuntos de nodos del conjunto, $D = (V, E)$, siendo éste un DAG. Se dice que Z *d*-separa X e Y , y se denota como $X \perp_D Y \mid Z$, si y solo si, para cualquier camino no dirigido entre un nodo de X y un nodo de Y existe un nodo intermedio V tal que

1. Existe una conexión convergente con V , siendo éste el nodo al que convergen las aristas, tal que ni V ni los descendientes de V pertenecen a Z .
2. Bien la conexión es en serie y V es un nodo intermedio; bien la conexión es divergente siendo V el padre, y V pertenece a Z .

Definición 1.16 (*d*-separación, (Lauritzen) Sean X , Y y Z tres conjuntos disjuntos de nodos del conjunto, $D = (V, E)$, siendo éste un DAG. Se dice que Z *d*-separa X e Y , y se denota como $X \perp_D Y \mid Z$, si y solo si Z separa X e Y en el grafo moral del menor subconjunto ancestral que contenga a los nodos de X , Y y Z , dado por $\mathcal{D}_{an(XUYUZ)}^m$.

En segundo lugar, dado un modelo gráfico probabilístico como puede ser un DAG en el que, como hemos dicho anteriormente, los nodos representan variables aleatorias y los arcos la estructura de dependencia entre ellas, podemos intuir el concepto de **evidencia** al conocimiento del valor exacto que toma una de las variables aleatorias, ya que este valor condiciona la incertidumbre del resto de variables, y por tanto, a la información de la red completa. En función del conocimiento de los valores exactos de una o más variables, la información probabilística de la red completa se va actualizando, de tal manera que cada actualización puede suponer una mejora respecto al estado precedente.

Al proceso de actualización de la probabilidad de la red a partir de la información de alguna de las variables del problema se le denomina propagación de la evidencia.

La característica de propagación es importante en las Redes Bayesianas, como un tipo de modelos gráficos probabilísticos que son, considerando que éstas pueden incluir variables observables y no observables (continuas o discretas), así como los

parámetros de los modelos de inferencia bayesiana, de tal forma que sea posible conformar distribuciones iniciales sobre las variables que no tienen padres en el DAG, así como distribuciones condicionadas de cada variable por sus padres, para el resto de variables.

Este hecho es fundamental, como se verá en los siguientes capítulos, al ser pieza clave en la demostración de que la aproximación bayesiana supone una mejora en los modelos de transporte.

Capítulo 2

Redes Bayesianas (RB)

2.1. Antecedentes

Si bien los modelos usuales matemáticos tienen una larga trayectoria y una base sólida a la hora de encontrar soluciones al problema de la elección discreta, existen otro tipo de aproximaciones diferentes a las frecuentistas que pueden dar lugar a mejores resultados.

Este es el caso de las aproximaciones bayesianas, que en este contexto concreto de los modelos de transporte se referirán como redes bayesianas.

Las redes bayesianas surgen en la década de los 80, derivadas de las investigaciones que desde los 70 se venían haciendo en inteligencia artificial (IA) con los sistemas expertos, programas capaces de simular e incluso sustituir en algunas ocasiones a los razonamientos humanos.

El término de redes bayesianas”se le atribuye en estos años, en concreto, en el año 1985 a Judea Pearl¹, (Pearl, J. 1985) para hacer hincapié en tres aspectos fundamentales: el carácter a menudo subjetivo de la información de entrada; la dependencia de acondicionamiento de Bayes como base

¹ J. Pearl, informático y filósofo, conocido por desarrollar la aproximación probabilística a la IA, en particular utilizando las Redes Bayesianas, y la formalización del razonamiento causal en las ciencias empíricas. Desde 1970 trabaja en UCLA, donde actualmente es profesor en Ciencias de la Computación y Estadística y director del Laboratorio de Sistemas Cognitivos. Recientemente, en 2011 ha recibido el premio Turing por sus contribuciones fundamentales a la IA a través del desarrollo del cálculo de probabilidades y del razonamiento causal.

para la actualización de la información; la distinción entre los modos causales y probatorios de razonamiento, lo que subraya Thomas Bayes en un documento publicado póstumamente en 1764 (Bayes, T.; Price, Mr. 1764).

Pearl se basa en la utilización de la probabilidad subjetiva que ha sido introducida por Thomas Bayes en su célebre ensayo de 1964. Esto le permite a Pearl basarse en la fórmula de Bayes para introducir las “redes bayesianas”, como un procedimiento para actualizar la información.

A finales de 1980 los textos seminales *Razonamiento Probabilístico en Sistemas Inteligentes* (Pearl, J. 1988) y *Razonamiento Probabilístico en Sistemas Expertos* (Neapolitan, Richard E. 1989) resumen las propiedades de las redes Bayesianas y ayudan a establecer a las mismas como un campo de estudio.

Siguiendo esta línea de investigación, durante la década de los 80 y 90, se puso de manifiesto que la IA no debía únicamente imitar el comportamiento racional humano, si no colaborar con éstos en las tomas de decisión a través de sinergías. Es decir, transmitir conocimiento del proceso lógico seguido para resolver un problema y obtener una solución.

Al respecto, en el año 1993, Clancey dirige una cita a los autores de la revista Knowledge Acquisition, quienes deciden publicarla, en la que indica: “La cuestión clave no es la IA, sino cómo mejorar la inteligencia natural con la ayuda de los sistemas basados en conocimiento.”²

2.2. Definición y propiedades

Dicho esto como antecedente contextual, para explicar la importancia que en las últimas décadas están teniendo las RB, expliquemos que una RB es un modelo gráfico probabilístico, un grafo acíclico dirigido (DAG) que representa:

1. cualitativamente, un conjunto de variables aleatorias (nodos³) y sus **interdependencias** condicionales probabilísticas (codificado en sus arcos) y;

² “The key issue is not artificial intelligence, but how to extend natural intelligence through knowledge based systems.”

³ Los nodos pueden representar cualquier variable: un parámetro medido, una variable latente o una hipótesis.

2. cuantitativamente, las **distribuciones de probabilidad condicionadas** de cada nodo dado sus padres.

Por tanto, tal y como se ha indicado anteriormente, las variables aleatorias pueden ser discretas o absolutamente continuas, por lo que las distribuciones de probabilidad asociadas serán funciones de masa y funciones de densidad respectivamente.

Para la especificación de la información cualitativa de la Red Bayesiana, se utiliza un DAG, que denotamos $D = (V, E)$, donde cada uno de los nodos de D representa las variables del problema $X = \{X_1, \dots, X_n\}$, siendo por tanto $V = \{X_1, \dots, X_n\}$; y las aristas dirigidas que están en E muestran las relaciones de tipo causal, siendo él o los nodos padre, la causa, y él o los nodos hijos, el efecto.

Además, la especificación de la información cuantitativa viene dada por un conjunto de distribuciones de probabilidad condicionada

$$P = \{p(x_1 | \text{pa}(X_1)), \dots, p(x_n | \text{pa}(X_n))\},$$

de forma que para cada variable X_i X se tendrá la distribución de probabilidad condicionada de X_i dada la ocurrencia de sus padres $\text{pa}(X_i)$ en el grafo D , denotada por $p(x_i | \text{pa}(X_i))$.

En resumen, formalmente, uniendo los dos conceptos, una RB está formada por el par (G, P) , donde G es un DAG formado por un nodo para cada variable aleatoria de $X = \{X_1, \dots, X_n\}$ y arcos que representan la estructura de dependencia probabilística entre ellas, $P = \{p(x_1 | \text{pa}(x_1)), \dots, p(x_n | \text{pa}(x_n))\}$ es un conjunto de n distribuciones de probabilidad condicionadas, y $\text{pa}(x_i)$ es el conjunto de padres del nodo X_i en G .

Es decir, un DAG es una RB respecto a un conjunto de variables, si el conjunto de la distribución de probabilidad de las variables nodo puede ser escrito como el producto de la distribución local de cada nodo y sus padres como la siguiente factorización cuya demostración puede verse en Jensen (2001):

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | \text{padres}(X_i)) \quad (1)$$

Si hay un arco del nodo A a otro nodo B , A es llamado padre de B y B es un hijo de A . El conjunto de nodos padre de un nodo x_i se denota como $\text{padres}(x_i)$.

En esta última ecuación anterior se tiene una distribución de probabilidad condicionada por cada variable aleatoria. O lo que es lo mismo, cada nodo X_i es una variable condicionada por sus padres, lo que establece una relación directa entre la parte cualitativa y la parte cuantitativa de la red, ya que es el DAG el que permite determinar las distribuciones de probabilidad condicionadas que se consideren en la factorización de la distribución de probabilidad conjunta. Es decir, a un DAG le corresponde una factorización de la distribución de probabilidad conjunta de una RB.

A partir de esta propiedad, existen algoritmos eficientes que realizan inferencia y aprendizaje en RBs como Neapolitan (2004), Castillo y otros (1997).

Entre las principales características de las RB se debe destacar que cumplen las propiedades de *d-separación*, que determinan estructuras de independencia (y dependencia) condicionada. Esto es, a partir del concepto de *evidencia* o valor exacto conocido que toma una de las variables (nodos), tal que al introducir esta información en la red, afecta a la incertidumbre del resto de las variables, se estudia cómo se traspa la información de dicha evidencia a lo largo de una red, es decir, los criterios de separación que se cumplen en el grafo; las relaciones de independencia o dependencia condicionada entre las variables, que dependerán del tipo de conexión. Así se pueden considerar:

1. conexión **serial**: se transmite información de la evidencia, excepto que dicha información esté contenida en el nodo intermedio, (A y C *d-separados* dado B)
2. conexión **divergente**, la información fluye a través de la red, excepto si la evidencia se encuentra en el nodo padre, ya que queda bloqueada la comunicación entre los nodos hijos, (B y C *d-separados* dado A)
3. conexión **convergente**, la información puede ser transmitida a través de la red sólo si se tiene la evidencia sobre el nodo hijo, o un descendiente de éste, (A y B *d-conectados* dado C)

Otra propiedad importante es la de Markov tal que si y sólo si, cada nodo X_i es condicionalmente independiente de sus no descendientes, $nd(x_i)$, dado sus padres, $pa(x_i)$ (Korb y Nicholson (2004)). Es decir,

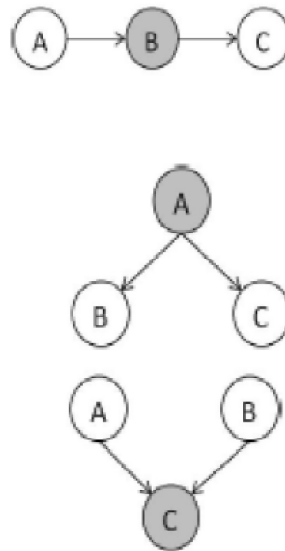


Figura 2.1: De arriba a abajo: conexión serial, conexión convergente y conexión divergente

$$P(X_i | pa(X_i), nd(X_i)) = P(X_i | pa(X_i)) \quad (2)$$

En la Figura 2.2 se muestra un ejemplo de red con la propiedad de Markov:

Esta propiedad, permite definir una RB de forma similar a la ecuación (1), y pueden demostrar los siguientes teoremas:

Teorema 2.1: *Todo par (G, P) que cumple la propiedad de Markov, constituye una red Bayesiana.*

Teorema 2.2: *Toda red Bayesiana formada por el par (G, P) , cumple la propiedad de Markov.*

Conclusión, considerando recíprocos estos dos teoremas, se puede concluir que toda RB definida por el par (G, P) , cumple dos propiedades equivalentes: la factorización presentada en (1) y la de Markov presentada en (2).

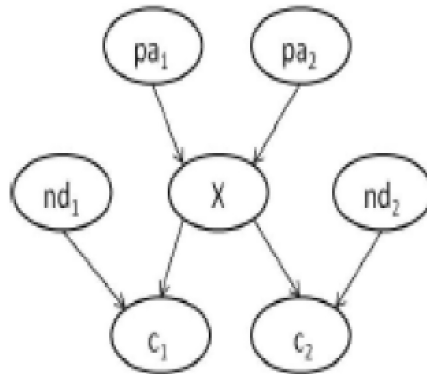


Figura 2.2: Red con propiedad de Markov

2.3. Tipología y aplicaciones

Se debe señalar que existen distintos tipos de RB, dependiendo de la tipología de la variable aleatoria X considerada en ella, discreta y/o continua, y que estas redes facilitan el desarrollo de distribuciones de probabilidad complicadas.

En una RBD todas las variables aleatorias de $X = \{X_1, \dots, X_n\}$ son discretas, lo que además implica que sus posibles estados son finitos; si además son binarias, respondiendo a los procesos de Bernoulli, la red se denota como Red Bayesiana Multinomial.

2.3.1. Redes bayesianas discretas (RBD)

Dado que parte de la presente memoria se ha realizado durante la pandemia producida por el virus COVID19 (SARS-CoV-2), se incluye un ejemplo relacionado. En la siguiente imagen se muestra una RB Discreta, asociada a los eventos dicotómicos (SÍ=1, No=0) de tener: Ictus, COVID19 y síntomas neurológicos, como cefalea, anosmia o dolores musculares, además de otros con afectación del sistema nervioso central y periférico cuya frecuencia y alcance están aún por determinar, según lo publicado en el segundo trimestre de 2020 por la Sociedad Española de Neurología (SEN) en el “Manual Covid-19 para el neurólogo general”.

1. tener un Ictus está relacionado con tener síntomas neurológicos

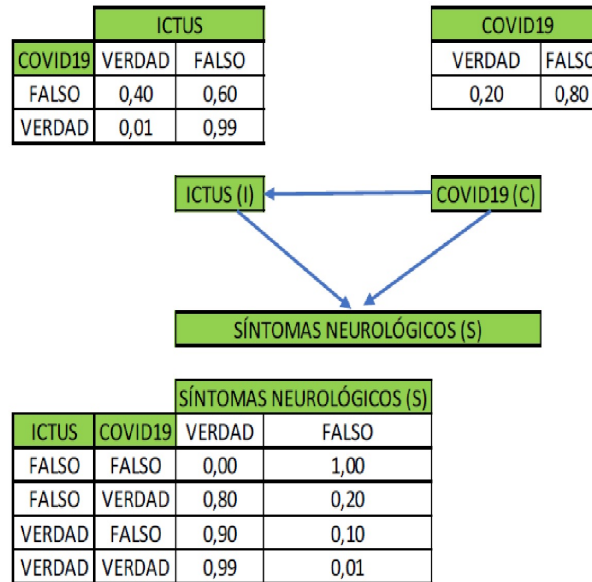


Figura 2.3: Red Bayesiana Discreta

2. tener COVID19 puede estar relacionado con tener un Ictus
3. tener COVID19 puede estar relacionado con tener síntomas neurológicos

De tal manera que se podrían configurar las tablas de probabilidad de la Figura 2.3 asociadas a esta RB y a partir de ella calcular la distribución de probabilidad conjunta como:

$$P(C, S, I) = P(C) P(I|C) P(S|I, C).$$

2.3.2. Redes bayesianas gaussianas (RBG)

Estas redes incorporan variables de tipo continuas, donde se asume que las variables tienen distribución Normal y por tanto la relación entre ellas es lineal; una familia Potencial Exponencial Multivariante (MEP), distribución que ha sido ampliamente desarrollada por este grupo de investigación de Métodos Bayesianos; y también puede usarse una aproximación de las densidades condicionadas usando distribuciones de sumas Gaussianas ponderadas.

Derivado de lo anterior, por tanto, en las RBG la distribución conjunta de las variables del problema $X = \{X_1, \dots, X_n\}$ es normal multivariante $N(\mu, \Sigma)$, tal que la función de densidad conjunta es:

$$f(x) = (2\pi)^{-n/2} |\Sigma|^{-1/2} \exp \left[-\frac{1}{2} (x - \mu)^t \Sigma^{-1} (x - \mu) \right],$$

siendo

1. vector de medias de dimensión n
2. Σ matriz de covarianzas, definida positiva de dimensión $n \times n$
3. $|\Sigma|$ determinante de la matriz de covarianzas
4. $(x - \mu)^t$ el vector traspuesto de $x - \mu$.

Además, al ser una RB se cumple lo indicado en el apartado anterior en cuanto a su factorización:

$$P(X_1, \dots, X_n) = \prod_{i=1}^n P(X_i | \text{padres}(X_i))$$

y aplicando las propiedades de la distribución Normal, tenemos:

$$f(X_i | \text{padres}(X_i)) \sim N \left(\mu_i + \sum_{j=1}^{j-1} \beta_{ji} (x_j - \mu_j), \nu_i \right),$$

siendo

- a) β_{ji} el coeficiente de regresión de X_j en la regresión de X_i condicionada a sus padres
- b) ν_i la varianza condicionada de X_i dado sus padres, tal que,

$$\nu_i = \Sigma_i - \Sigma_{i \text{ pa}(X_i)} \Sigma_{\text{pa}(X_i)}^{-1} \Sigma_{\text{pa}(X_i)}^t$$

- c) Σ_i la varianza marginal de X_i

- d) $\sum_{ipa(X_i)}$ el vector de covarianzas entre X_i y las variables del conjunto $padres(X_i)$
- e) $\sum_{pa(X_i)}$ la matriz de covarianzas del conjunto $padres(X_i)$

Hay que tener en cuenta que el coeficiente de regresión β_{ji} indica la intensidad de la relación entre X_i y X_j . Esto significa que si $\beta_{ji} = 0$, tal que $j < i$, entonces X_j no es padre de X_i , y por tanto no existe una arista dirigida entre ellas.

Históricamente dos de los ejemplos más empleados para las RBG han sido los de Enrique Castillo Ron, en cuanto a la aplicación de la Estadística (Castillo et al., 2003): 1) estudio del caudal de un río; 2) evaluación del nivel de daño de una viga de hormigón.

Otro ejemplo más reciente sobre el estudio de la botánica es el incluido en Scutari, M. y Denis, J.B. (2015), en el que se analiza el crecimiento de una planta considerando:

1. Potencial del entorno, $E \sim N(50 \ 10^2)$
2. Potencial genético, $G \sim N(50 \ 10^2)$
3. Organismos vegetales, $V|G, E \sim N(-10.35534+0.5G+0.7711E, 5^2)$
4. Número de semillas, $N|V \sim N(45+0.1V, 9.949874^2)$
5. Peso medio de las semillas, $W|V \sim N(15+0.7V, 7.141428^2)$
6. Masa de grano cosechada, $C|N \sim N(0.3N+0.3W, 6.25^2)$

Si bien, un ejemplo más relacionado con el objeto de la presente memoria es el incluido en Weidl, G., Madsen, A. L., Wang, S., Kasper, D. y Karlsen, M. (2018), en el que se usan RBG Dinámica para entender el comportamiento de los patrones de tráfico y el reconocimiento temprano de las maniobras en las carreteras por parte de los vehículos, analizando el comportamiento observado de los vehículos y el espacio libre disponible en el carril objetivo al realizar desplazamientos laterales entre carriles. Estos sistemas se utilizan en la conducción autónoma de vehículos y es otra de las líneas de investigación derivadas de esta investigación.

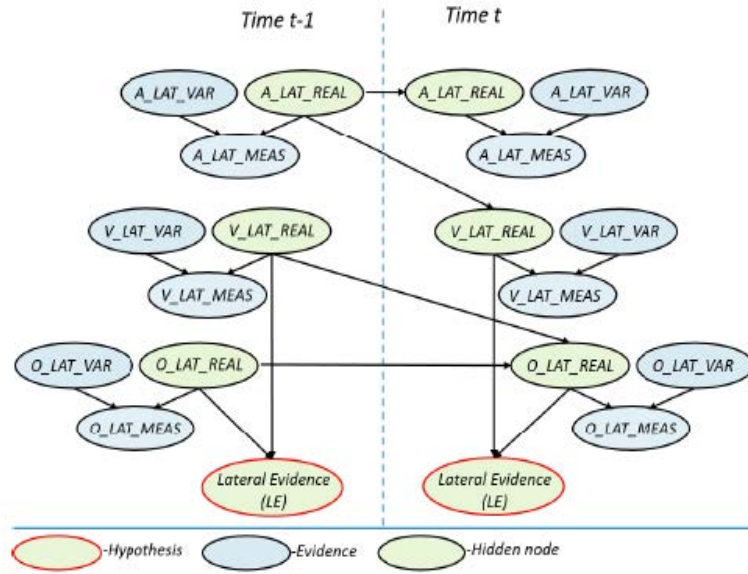


Figura 2.4: Red Bayesiana Gaussiana (desplazamientos vehiculares laterales con aceleración)

Entre los intervalos de tiempo $t - 1$ y t se asumen condicionales lineales gaussianos y se derivan de modelos físicos, como:

1. O LAT REAL (t) (valor esperado del desplazamiento lateral)
2. V LAT REAL (t) (valor esperado de la velocidad lateral)
3. A LAT REAL (t) (valor esperado de aceleración lateral)

Cuando O LAT REAL (t) es constante creciente y V LAT REAL (t) es alto o creciente (requiere un valor positivo A LAT REAL (t)), su combinación claramente indica que el vehículo está saliendo de su carril.

Y las variables V LAT REAL(t) y A LAT REAL(t) se definen tal que:

1. LAT REAL(t) con distribución

$$N(OLATREAL(t - 1) + VLATREAL(t - 1)\Delta_t, \sigma_{O_LAT(t)}^2)$$

2. V LAT REAL(t) con distribución

$$N(VLATREAL(t-1) + ALATREAL(t-1)\Delta_t, \sigma_{V_LAT(t)}^2)$$

3. A LAT REAL(t) $\sim N(ALATREAL(t-1)\Delta_t, \sigma_{V_LAT(t)}^2)$

2.3.3. Redes bayesianas mixtas (RBM)

Las redes bayesianas mixtas o híbridas contienen en el que el conjunto de variables considerado $X = (X_1, X_2, \dots, X_n)$ variables de tipo continuo y discreto, tal que, las variables discretas tienen número finito de estados (exclusivos y excluyentes), precediendo a las continuas en el grafo; y las variables continuas son Gaussianas, bien con distribución condicionada lineal Gaussiana de sus padres discretos, bien con la media de la distribución condicionada Gaussiana lineal del estado de sus padres continuos. No es posible la existencia de padres continuos e hijos discretos. Información al respecto de este tipo de redes puede encontrarse en Cowell, et al. (1999).

Calcular la distribución conjunta en este tipo de redes e inferir es más complejo, ya que la propagación de la evidencia se hace más complejo.

Dicho esto, es interesante la utilización de mezclas de exponenciales truncadas (MTE) por parte de Moral et al. (2001), para representar la distribución de las variables de la red, discretas y continuas. De esta forma, las variables discretas pueden tener padres continuos en la red, lo que significa bajar el nivel de las restricciones anteriormente indicadas para las RBM habitualmente definidas. Información al respecto del proceso de inferencia y aprendizaje en RBM puede encontrarse en Cobb et al. (2007)

En cuanto a la representación gráfica del DAG de una RBM, para distinguir entre variables discretas y continuas, se considera una doble circunferencia (o doble elipse) en las variables continuas.

El conjunto de nodos $V = \{X_1, \dots, X_n\}$ se particiona en variables discretas (Y) o variables continuas (Z), siendo $V = Y \cup Z$, de tal manera que en un modelo condicionado Gaussiano, la distribución de una variable continua, tal que sus padres son discretos, será una distribución Gaussiana Multivariante:

$$\mathbf{Z} \mid \mathbf{Y} = \mathbf{y} \sim N(\mathbf{y}, \Sigma(\mathbf{y}))$$

con Σ definida positiva.

Definición 2.1 Una variable aleatoria mixta $\mathbf{X} = (\mathbf{Y}, \mathbf{Z})$ tiene **distribución condicionada Gaussiana** si la función de densidad conjunta es:

$$f(\mathbf{x}) = f(\mathbf{y}, \mathbf{z}) = \chi(\mathbf{y}) \exp\{g(\mathbf{y}) + \mathbf{h}(\mathbf{y})^t \mathbf{z} - \mathbf{z}^t K(\mathbf{y}) \mathbf{z} / 2\}$$

tal que:

1. $\chi(\mathbf{y}) \in \{0, 1\}$ representa si f es positiva en \mathbf{y}
2. g , función real
3. h , vector
4. K , matriz

2.3.4. Aplicaciones

En cuanto a las aplicaciones, de forma resumida, y por no retrotraernos demasiado atrás en el tiempo, los modelos gráficos probabilísticos se han usado para realizar técnicas de diagnóstico médico (Jensen, 2001), en sistemas expertos (Cowell, et al., 1999) y de planificación y control (Dean, et al., 1991; Chan, et al., 1992), además de sistemas dinámicos y de planificación de infraestructuras, como es este caso.

En el caso concreto que nos ocupa de las RB en la modelización de transporte, caben destacar en los últimos años, por citar algunos, los siguientes trabajos realizados:

- Aplicación de las redes bayesianas a la nueva movilidad mundial, como puede verse en Gómez-Villegas (2019), trabajo mostrado en el “XXXVIII Congreso Nacional de Estadística e Investigación Operativa - SEIO 2019” en Alcoi.

- Comparación de resultados en la elección modal del transporte de mercancías (De Gregorio- Vicente et al. (2017): “Aproximación Bayesiana aplicada al Reparto Modal en Modelos de Transporte de Mercancías. Caso Práctico: Corredor Ferroviario Bioceánico Central)

- Estimación de matrices origen-destino de viajes (matrices OD), a partir de la información de placas de matrículas en flujos vehiculares, o a partir de flujos parciales de tráfico dada una red urbana (Nogal Macho, M. (2011)), considerando ejemplos aplicados la red viaria de Nguyen-Dupuis, a la de Ciudad Real y a la del Estado de Vermont. Hay que indicar que dicho trabajo de Tesis Doctoral consiguió el primer Premio Internacional Abertis en el año 2012;

- Análisis de decisiones en diferentes escenarios de transporte, que puede verse en Holland, A. (2003);

- Construcción de modelos de predicción de flujos de tráfico en los arcos de una red dada, para estimar matrices OD, así como para encontrar la mejor ubicación para los puntos de conteo de tráfico con el objetivo de que las Administraciones (locales, regionales y nacionales) consigan una mejor gestión de la movilidad en términos globales, puede verse en Sánchez-Cambronero García-Moreno, S. (2008).

Todos estos trabajos han puesto de manifiesto la potencia de las RB en este tipo de investigaciones asociadas a la modelización y planificación de redes de transporte, considerando una cantidad importante de variables, discretas y continuas.

Si bien estos trabajos siempre han ido dirigidos hacia la estimación de flujos, marcando una línea clara de investigación, la presente memoria se dirige más hacia la utilización de las RB en cuanto a la optimización de los resultados obtenidos en la etapa del reparto modal de los modelos clásicos de transporte de cuatro etapas, que se enumeran a continuación y cuyo detalle se expone en el siguiente capítulo: primera etapa, Generación-Atracción (o Consumo); segunda etapa, Distribución; tercera etapa, Reparto modal; y cuarta etapa, Asignación a red.

2.4. Redes Bayesianas Gaussianas (RBG)

Vistos los aspectos fundamentales de las RB, definiciones, características, tipologías y aplicaciones, nos centraremos en los puntos clave de las RBG, al ser las más habitualmente utilizables, por las buenas propiedades derivadas de sus distribuciones Normales.

2.4.1. Construcción de una RBG

Los pasos elementales que considero a la hora de construir una RBG se pueden considerar una mezcla entre la metodología de 7 etapas propuesta por Bromley (2005) y las 3 etapas definida en otros trabajos similares (Estructura + Parámetros + Evidencia):

1. Identificar el marco de estudio y lo que se pretende resolver.
2. Conocer con qué inputs de entrada (variables y valores) fiables se puede contar. Esto es relevante ya que, dependiendo de la calidad de las variables, así será la calidad del modelo y de los resultados obtenidos. Esto lo resumen las metodologías anglosajonas al respecto de la construcción de modelos al mencionar: “garbage in, garbage out”.
3. Conocer las relaciones de dependencia entre las variables para poder construir el grafo asociado al modelo de red, lo cual implica la parte cualitativa de la red.
4. Calcular las distribuciones de probabilidad de cada nodo del DAG en cuanto a sus parámetros, lo cual implica la parte cuantitativa del grafo, en cuanto a las funciones de densidad o condicionadas.
5. Analizar la propagación de la evidencia en los nodos de la red, una vez definida la parte cualitativa y cuantitativa, en cuanto a saber cómo cambian las distribuciones de probabilidad a lo largo de la red cuando los valores de algunas de las variables son conocidos.

La construcción del DAG, a partir de las variables y valores disponibles, ésta puede llevarse a cabo mediante un proceso:

1. Manual: grupo de trabajo formado por expertos en cada ámbito, bien a partir de las distribuciones condicionadas, bien a partir de la distribución conjunta; normalmente, para agilizar la construcción, primero se definen los parámetros de las distribuciones condicionadas y después la distribución conjunta, partiendo de la matriz de covarianzas Σ . A partir de los parámetros definidos en

las densidades condicionadas, se puede obtener dicha matriz Σ y por tanto, formular la distribución conjunta de la RBG.

2. Automático: algoritmos de aprendizaje a partir de una base de datos dada, tal que se construya el DAG y las distribuciones de probabilidad.

Construcción manual con las distribuciones condicionadas. Los parámetros requeridos se definen en base a las distribuciones condicionadas de la expresión anteriormente indicada:

$$f(X_i | \text{padres}(X_i)) \sim N\left(\mu_i + \sum_{j=1}^{j-1} \beta_{ji}(x_j - \mu_j), \nu_i\right),$$

siendo

1. μ : el vector de medias con dimensión $n \times 1$, donde cada μ_i es la media de la variable X_i para $i = 1, \dots, n$.
2. los coeficientes de regresión β_{ji} de X_j en la regresión de X_i sobre $\text{pa}(X_i)$, para $i, j = 1, \dots, n$.

Estos dos conjuntos de parámetros son necesarios para determinar las medias condicionadas $E(X_i | \text{padres}(X_i))$.

3. las varianzas condicionadas ν_i de cada X_i dado sus padres, con $i = 1, \dots, n$.

Construcción manual con la distribución conjunta. Se parte de los parámetros de la distribución conjunta del DAG:

1. μ : el vector de medias con dimensión $n \times 1$, donde cada μ_i es la media de la variable X_i para $i = 1, \dots, n$.
2. Σ : la matriz (definida positiva) de covarianzas de dimensión $n \times n$, donde σ_{ii} es la varianza de X_i y σ_{ij} es la covarianza entre X_i y X_j , para $i, j = 1, \dots, n$.

Construcción automática mediante algoritmos. Como se ha indicado anteriormente, los algoritmos utilizados son capaces de construir el DAG y las distribuciones de probabilidad a partir de una base de datos inicial. La clave está en que dicha base contenga información suficiente y consistente para conseguir que los algoritmos funcionen de manera eficiente.

Se distinguen diferentes tipos de algoritmos:

1. basados en medidas: utilizando búsquedas heurísticas se fija una medida posible a la RB a partir de la base de datos inicial, tal que se maximice esta medida.
2. basados en restricciones: el objetivo es formar un DAG que verifique las reglas de d-separación aplicando test de independiencia condicionada consistente con las propiedades de Markov.
3. Mixtos o híbridos: combinan algoritmos basados en restricciones, que acoten el espacio de búsqueda; y algoritmos basados en medidas, tal que la medición del DAG sea máxima.

El caso de estudio presentado en esta memoria precisamente utiliza el algoritmo mixto *max – min hill – climbing* (mmhc) de Tsamardinos et al. (2006) para calibrar el DAG de una RB construida con una base de datos inicial, cuyos datos son oficiales y han sido depurados y validados previamente para garantizar la consistencia de la información.

Las restricciones del algoritmo son *max-min parents & children* (mmpc), también de Tsamardinos et al. (2003), para conformar la estructura de la red; y las medidas *hill-climbing* (hc), de Russell and Norvig (2009), para concretar el sentido de los enlaces.

2.4.2. Independencia e Independencia condicionada

Si bien son conocidos los siguientes resultados de la distribución Normal Multivariante mostrados en este apartado, son igualmente importantes a la hora de relacionar la independencia condicionada con la matriz de covarianzas Σ . Las demostraciones de estas proposiciones se encuentran en Lauritzen (1996) o Anderson (2003), respectivamente, entre otros autores.

Proposición 2.1 Sea \mathbf{X} con distribución Normal Multivariante $N(\mu, \Sigma)$ tal que \mathbf{X} se particiona como $\mathbf{X} = \{\mathbf{X1}, \mathbf{X2}\}$ con μ y Σ dados por

$$\mu = (\mu_1, \mu_2) \text{ y } \Sigma = \begin{pmatrix} \Sigma_{11} & \Sigma_{12} \\ \Sigma_{21} & \Sigma_{22} \end{pmatrix}.$$

Entonces, se cumple que:

1. $\mathbf{X1} \sim N(\mu_1, \Sigma_{11})$ y $\mathbf{X2} \sim N(\mu_2, \Sigma_{22})$
2. $\mathbf{X1}$ y $\mathbf{X2}$ son independientes si y solo si $\Sigma_{12} = \Sigma_{21} = \mathbf{0}$

Teniendo en cuenta que la matriz Σ es definida positiva, entonces su inversa existe y la matriz de precisión se puede obtener como $W = \Sigma^{-1}$. Entonces, llegamos a la siguiente proposición:

Proposición 2.2 Sea $X \sim$ Normal Multivariante $N(\mu, \Sigma)$, si Σ es una matriz regular e invertible, entonces las variables X_i y X_j son condicionalmente independientes dado el resto de las variables de X , si y sólo si, $W_{ij} = 0$ (valor de la posición (i,j) de la matriz de precisión K de la distribución es cero).

2.4.3. Propagación de la evidencia en RBG

En apartados anteriores se ha explicado en qué consiste la propagación de la evidencia partiendo de la información conocida de una o algunas variables, y cómo se pueden actualizar las distribuciones de probabilidad del resto de las variables en base a esta información.

Históricamente podemos identificar algoritmos de propagación desarrollados por nombrar a Normand, et al. (1992), Lauritzen (1992), Lauritzen, et al. (2001), Cowell (2005), si bien también hay que tener en cuenta el desarrollado por Castillo et al. (1997), que calcula recursivamente en cada iteración que se introduce información de la variable evidencial, la densidad condicionada de una distribución Normal, actualizando así las probabilidades no evidenciales de la red, en tiempo lineal, dada la evidencia.

Entonces, partiendo del conjunto de variables no evidenciales Y y de variables evidenciales $E = \{X_e\}$, entonces, X es una partición tal que $X = (Y, E)$, tal que la

distribución condicionada de Y dado $E = e$ es una distribución Normal Multivariante con parámetros:

1. $\mu^Y | \mathbf{E} = \mathbf{e} = \mu_Y + \Sigma_{YE} \Sigma_{EE}^{-1} (\mathbf{e} - \mu_E)$
2. $\Sigma^Y | \mathbf{E} = \mathbf{e} = \Sigma_{YY} - \Sigma_{YE} \Sigma_{EE}^{-1} \Sigma_{EY}$

Por otro lado, partiendo de la densidad marginal a posteriori de una única variable respuesta, $X_i \in Y$, y una variable evidencial E , entonces, al realizar la propagación tenemos la siguiente ecuación cuyos parámetros son anteriores a la propagación de la evidencia:

$$X_i | E = e \sim N(\mu_i^{Y|E=e}, \sigma_{ii}^{Y|E=e}) = N(\mu_i + \frac{\sigma_{ie}}{\sigma_{ee}} (e - \mu_e), \sigma_{ii} - \frac{\sigma_{ie}^2}{\sigma_{ee}}),$$

siendo

1. μ_i y e las medias de X_i y E_i
2. σ_{ii} y σ_{ee} las varianzas de X_i y E
3. σ_{ie} la covarianza entre X_i y E .

Tal y como se ha indicado anteriormente, la propagación de la evidencia es un proceso recursivo tal que la variable evidencial se actualiza iteración a iteración. Por tanto, existe una relación lineal entre la cantidad de operaciones requeridas para actualizar la distribución de probabilidad de las variables de Y , y el número de variables de X .

2.4.4. Normalidad y No Normalidad

Todo lo visto anteriormente es la base teórica que se considera imprescindible a la hora de afrontar un problema de redes. Sin embargo, la aplicación práctica de esta base no siempre es sencilla y para garantizar la fiabilidad y consistencia de los datos de partida es imprescindible validar los supuestos iniciales para poder construir una RBG con distribución de probabilidad conjunta, asociada a las variables de partida $X = (X_1, X_2, \dots, X_n)$, Normal Multivariante.

Para la validación del supuesto de Normalidad, el software R <https://cran.r-project.org/> dispone de una librería llamada *asbio* en la que está implementada el test Omnibus de Doornick-Hansen (2008), que además es más potente que el test de Shapiro-Wilks.

Este test es de gran ayuda desde varios puntos de vista, tomando de partida una base de datos con un número considerable de variables y registros:

1. si cada una de las variables, de forma independiente, no se ajusta a una distribución Normal Univariante, entonces la matriz no se ajusta a una distribución Normal Multivariante.
2. si todas las variables se ajustan a una distribución Normal Univariante, esto no implica el hecho de que la matriz total se ajuste a una distribución Normal Multivariante.
3. además, a nivel gráfico, revisar un gran número de variables aleatorias en cuanto a su supuesto de distribución Normal resulta complejo, por solapamientos, por cambios de escala, etc.

2.4.5. Test de Doornick-Hansen

La base teórica del estadístico Doornick-Hansen es la siguiente: a) Caso univariante: toma de referencia el estadístico de Bowman y Shenton (1975), basado en el cálculo de kurtosis, y asimetría multivariante, que si bien son incorrelados, no cumplen con ser independientes, y además, la kurtosis de la muestra se acerca paulsadamente a la condición de Normalidad. La solución para todo esto deriva de las investigaciones realizadas por Shenton y Bowman (1977), de tal forma que se establece para la kurtosis una función de tipo gamma, tal que se cumpla que $kurtosis > asimétrica + 1$; b) Caso multivariante: se utiliza una transformación tal que aproxima una distribución multivariante a distribuciones Normales (estándar) independientes. Una vez hecho, posteriormente de forma univariante se obtienen kurtosis y simetría. Así, el test resultante obtiene vectores de kurtosis y simetría, en lugar de valores únicos.

El test se define tal que:

$$DH = Z_1^t Z_1 + Z_2^t Z_2$$

Donde las variables Z_1 and Z_2 son aproximaciones Normal transformadas considerando la asimetría (S) y la kurtosis (K) de la muestra.

La asimetría univariante (S), se transforma en una variable z_1 aproximadamente Normal, tal y como se indica en D'Ágostino (1970):

$$z_1 = \delta \ln (y + \sqrt{1 + y^2}),$$

con:

$$\beta = \frac{3(n^2 + 27n - 70)(n + 1)(n + 3)}{(n - 2)(n + 5)(n + 7)(n + 9)}$$

$$w^2 = -1 + \sqrt{2(\beta - 1)}$$

$$\delta = (\ln \sqrt{w^2})^{1/2}$$

$$y = S \left[\frac{(w^2 - 1)(n + 1)(n + 3)}{12(n - 2)} \right]^{1/2}$$

A su vez, la kurtosis univariante (K) se transforma de una Gamma a una χ^2 y luego en una Normal estándar, z_2 , utilizando la transformada de Wilson & Hilferty (1931):

$$z_2 = \sqrt{9a} \left[\left(\frac{\chi}{2a} \right)^{\frac{1}{3}} - 1 + \frac{1}{9a} \right],$$

con:

$$\chi = 2f(K - 1 - S^2)$$

$$a = a + S^2c$$

$$f = \frac{(n + 5)(n + 7)(n^3 + 37n^2 + 11n - 313)}{12d}$$

$$c = \frac{(n - 7)(n + 5)(n + 7)(n^2 + 2n - 5)}{6d}$$

$$a = \frac{(n - 2)(n + 5)(n + 7)(n^2 + 27n - 70)}{6d}$$

$$\delta = (n - 3)(n + 1)(n^2 + 15n - 4)$$

En el caso k-variable, el estadístico $Z_1^t Z_1 + Z_2^t Z_2$ es aproximadamente una χ^2 con $2k$ grados de libertad.

Alcanzado este punto y a la vista de los posibles resultados del test multivariante de Normalidad surge la pregunta inevitable: cómo solucionar la situación en la que los datos disponibles de partida no pasan dicho test, sabiendo que no dispongo de más datos ni de más recursos? Si me baso en la construcción manual de la RBG y dispongo de un grupo de trabajo formado por expertos, se puede plantear utilizar otras variables aleatorias que sustituyan a las inicialmente propuestas. Pero, si me baso en la construcción automática de la RBG, mediante algoritmos, se incumplen los supuestos necesarios de partida. La única solución factible pasa por transformar, de alguna forma, la base de datos inicial tal que se cumpla la condición de Normalidad, pero sin modificar la correlación existente entre ellos. De lo contrario, obtendríamos resultados a partir de la transformación, que no serían ni fiables ni acordes a los datos iniciales. Y por ello, no es factible la transformación de Box-Cox.

2.4.6. Distribución Nonparanormal

La solución a lo indicado anteriormente está en la distribución *Nonparanormal* para transformar datos, a partir del ajuste de grafos no dirigidos propuesta por Lui et al. (2009), lo cual garantiza cumplir la condición de Normalidad para el caso multivariante. Esta transformación está incluida en el paquete *huge* (2012) de R.

En las siguientes páginas se muestran las principales definiciones, propiedades, lemas que justifican la base teórica de esta transformación, fundamental para la utilización de bases de datos en RBG.

Definición 2.1 *Se dice que el vector $X = (X_1, \dots, X_n)$ tiene distribución Nonparanormal (npn) si existen funciones $\{f_j\}_{j=1}^n$ tales que:*

$$Z \equiv f(\mathbf{x}) \sim N(\mu, \Sigma) \text{ con } f(\mathbf{x}) = (f_1(x_1), \dots, f_n(x_n))$$

La solución está en la distribución *Nonparanormal* para transformar datos, a partir del ajuste de grafos no dirigidos propuesta por Lui et al. (2009), lo cual garantiza

cumplir la condición de Normalidad para el caso multivariante. Esta transformación está incluida en el paquete *huge* (2012) de R.

En las siguientes páginas se muestran las principales definiciones, propiedades, lemas que justifican la base teórica de esta transformación, fundamental para la utilización de bases de datos en RBG.

Entonces $\mathbf{X} \sim N(\mu, \Sigma, f)$, siendo las funciones f_j monótonas y diferenciables, de tal forma que la función de densidad conjunta de X es:

$$p_{\mathbf{X}}(\mathbf{x}) = \left((2\pi)^{n/2} |\Sigma|^{1/2} \right)^{-1} \exp \left(-\frac{1}{2} (f(x) - \mu)^t \Sigma^{-1} (f(x) - \mu) \right) \prod_{j=1}^n |f'_j(x_j)|.$$

Las siguientes propiedades garantizan la estimación del grafo:

1. Si $\mathbf{X} \sim N(\mu, \Sigma, f)$ y cada f_j es diferenciable, entonces la configuración de independencia condicionada de la información de partida, considerando las variables iniciales, permanece constante con la matriz Σ^{-1} .
2. El grafo asociado a la matriz Σ^{-1} , en cuanto a la configuración las variables que han sido transformadas manteniendo el supuesto de Normalidad, es igual en relación a la configuración de las variables de partida no Normales, lo cual se puede comprobar en los ceros de la matriz Σ^{-1} .

Teorema de Sklar (1959). *Cualquier función de distribución acumulada conjunta se puede escribir a partir de las funciones de distribuciones acumuladas marginales como:*

$$F(x_1, \dots, x_n) = C\{F(x_1), \dots, F(x_n)\}, \text{ donde } C \text{ es una función llamada copula.}$$

Teniendo en cuenta la distribución Nonparanormal, entonces:

$$F(x_1, \dots, x_n) = \Phi_{\mu, \Sigma}(\Phi^{-1}(F_1(x_1), \dots, F_n(x_n))),$$

siendo

1. $\Phi_{\mu, \Sigma}$ función de distribución de una Normal Multivariante
2. Φ función de distribución de una Normal (estándar) Univariante.

Entonces, se consigue una copula Gaussiana con parametros μ y Σ , tal que:

$$C(u_1, \dots, u_n) = \Phi_{\mu, \Sigma}(\Phi^{-1}(u_1), \dots, \Phi^{-1}(u_n))$$

Derivado del anterior teorema llegamos al siguiente lema:

Lema. *La distribución Nonparanormal es una cópula Gaussiana en el caso que las f'_j sean monótonas y diferenciables.*

Por otro lado, teniendo en cuenta la estructura de la diagonal de la matriz Σ , y no la matriz de covarianzas completa, se puede asumir que la densidad de X , con $\mathbf{X} \sim N(\mu, \Sigma, f)$, permanece igual si cumple la condición de que las funciones f_j mantienen el valor de las medias y las varianzas, es decir:

Sea $F_j(x)$ la función de distribución acumulada de X_j , teniendo en cuenta que $f_j(X_j)$ se ajusta a una función de densidad Normal, entonces:

$$F_j(x_j) = P(X_j \leq x) = P(Z_j \leq f_j(x)) = \Phi\left(\frac{f_j(x) - \mu_j}{\sigma_j}\right),$$

y

$$f_j(x) = \mu_j + \sigma_j \Phi^{-1}(F_j(x))$$

Considerando la anterior ecuación se pueden definir los parámetros y funciones de distribución marginales a estimar, para lo que se considera las especificaciones recogidas en el trabajo de Liu et al. (2009) y Lafferty et al. (2012).

Por tanto, maximizar $\log W'$ sujeto a las restricciones correspondientes al conocimiento sobre los macro-estados, permite generar modelos para estimar los meso-estados más probables. En nuestro caso la matriz más probable T.

Capítulo 3

Modelos de transporte

3.1. Introducción

El análisis de inversiones en grandes infraestructuras de transporte requiere la estimación, con largos horizontes de análisis de distintos aspectos. En primer lugar, cómo los usuarios de dichas redes (sean pasajeros o transportistas de mercancías) harán uso de las mismas si escogen unas alternativas frente a otras.

Para ello hay que ver si un nuevo proyecto supone una mejora real para el conjunto de la población y de su sistema de transportes, de tal manera que compense los altos costes de inversión a los que se ha de hacer frente, verificando esto mediante un análisis coste-beneficio (ACB).

Por tanto, conviene ver el nivel de endeudamiento y financiación necesarios, planteando posibilidades de internacionalización de la deuda, o planteando soluciones mixtas de participación público-privada (PPP).

De cara a dar respuesta a estas cuestiones, los modelos de transporte proporcionan una herramienta útil para la toma de decisiones ya que proporcionan un marco estructurado en el cuál comparar alternativas de diseño y estimar cuotas de mercado en base a criterios técnicos. Además, un modelo de transporte facilita el análisis de resultados bajo diferentes escenarios, con los cuales evaluar cómo pueden afectar a la viabilidad de una propuesta diversos riesgos o fuentes de incertidumbre.



Figura 3.1: Modelo macroscópico de Estados Unidos de América (Fuente: Caliper Corp.)

3.2. Fundamentos previos

3.2.1. ¿Qué son y para qué se usan?

Para el estudio de los modelos de transporte se necesita hacer representaciones virtuales de una red física, formada por arcos y nodos que poseen unas determinadas características, en la que participan diferentes modos de transporte (aéreo, marítimo/fluviial, terrestre).

Se usan para la planificación del transporte en determinados ámbitos de estudio (barrios, ciudades, regiones, países, continentes): se planifican las infraestructuras (red de carreteras, puertos, ff.cc., aeropuertos, hidrovías, corredores marítimos,...) actuales y futuras; se planifican las unidades a gestionar (vehículos, personas, mercancías), es decir, la demanda existente y la demanda proyectada.

En las Figuras 3.1 y 3.2 se muestra algún ejemplo.

3.2.2. Objetivos principales

Se debe empezar por investigar el comportamiento actual de pasajeros y mercancías, para poder predecir ese comportamiento a futuro y proyectar las demandas. Esto lleva a la identificación y evaluación de problemas para pasar a la planificación de las actuaciones que deban llevarse a cabo.



Figura 3.2: Modelo macroscópico de Madrid (Fuente: Ayuntamiento de Madrid (Central de Movilidad))

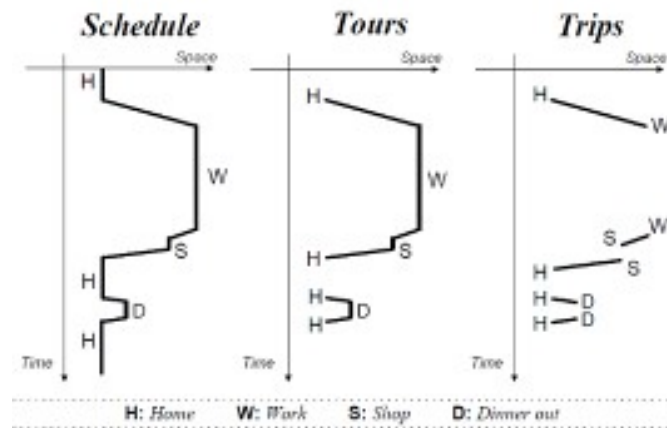


Figura 3.3: Desagregación de las diferentes etapas de viaje a lo largo de un día (Fuente: “Travel Demand Modelling” (Moshe Ben Akiva) – Massachusetts Institute of Technology (MIT))

3.2.3. Tipología y niveles de detalle

Existen dos tipologías principalmente de modelos macroscópicos: los que están basados en viajes, en los que la unidad es un viaje entre un origen y un destino y son denominados habitualmente como modelos clásicos de cuatro etapas (four-step algorithm) y cuyas etapas son las siguientes: generación-atracción, distribución, reparto modal y asignación a la red de transporte; los que están basados en actividades, en los que se analiza la cadena de viajes en un día completo derivada de realizar una serie de actividades (tour-based model).

En la siguiente figura se observa un ejemplo en el que se desagregan los viajes realizados durante un día, en diferentes actividades. Al realizar esta desagregación obtendré tantas matrices origen-destino como actividades individuales realizadas a lo largo del día, por lo que el ajuste de las matrices asociadas a cada actividad eleva la complejidad.

En cuanto a los niveles de detalle de los modelos de transporte, dependiendo del “zoom” que se haga de la red, existen los siguientes a estudiar: el macroscópico, el mesoscópico, el microscópico y el híbrido.

3.3. Modelo clásico de cuatro etapas

3.3.1. Introducción

Existe un gran número de metodologías que pretenden solucionar los problemas de transporte, tanto de mercancías como de pasajeros, y por lo tanto ayudar en las tomas de decisiones a efectos operativos y funcionales. Igualmente, en su evaluación económica, considerando que muchas de éstas siguen procesos similares en su desarrollo (Horn 2003). Inicialmente se define el ámbito geográfico, posteriormente se identifica la red de infraestructura que luego dará lugar a la red de arcos y nodos, se localizan los puntos de origen y destino de todos los posibles viajes (residenciales y productivos) que luego darán lugar a los centroides de la red (centros virtuales de gravedad), con darán lugar a las matrices origen-destino (OD) de los viajes. Estos orígenes y destinos son los fletes, en el caso de las mercancías, y los centros residenciales y de trabajo, en el caso de los viajeros. Además, han de ser identificados todos los posibles medios de transporte que se pueden emplear, y por cada uno de ellos, por lo menos, sus costes (o tarifas) y tiempos de viaje ya que son las razones principales en las que se basa el fletador o el viajero para elegir el medio más atractivo. Los tiempos de viaje, a su vez, están formados por los tiempos de acceso/dispersión desde los puntos de origen y destino

Además, considerando un transporte multimodal, como este caso, también hay que localizar y caracterizar las terminales intermodales (carretera-ferrocarril, ferrocarril-puerto, . . .).

Para ello se utiliza un modelo clásico de cuatro etapas de uso extendido en la planificación de transporte (Ortúzar, J.D., y Willumsen, L. G. (2011)) de mercancías y pasajeros, cuyas etapas se resumen a continuación:

1. Generación-Atracción. Una vez zonificada el ámbito de estudio y la red del modelo, se obtienen las toneladas de carga / número de viajeros generados/atraídos en cada TAZ (Traffic Analysis Zones) o ZT (Zona de Transporte), normalmente en formato vectorial, mediante el ID (identificador) de cada zona. De tal forma que se obtiene: ID Generación, ID Atracción, Toneladas (o Viajeros).
2. Distribución espacial. Se transforma la información de generación-atracción

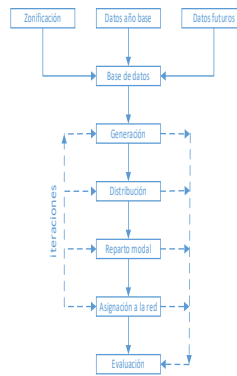


Figura 3.4: Modelo clásico de las cuatro etapas (Ortúzar and Willumsen 2011)

a origen-destino, dando como resultado un conjunto de matrices de origen-destino (matrices OD), por ID, cuyas celdas representan las toneladas de mercancías/número de viajeros. Explicar que, a su vez, estos resultados se pueden transformar a número de envíos (mercancías) y número de viajes (viajeros) en este punto o en el siguiente.

3. Reparto o elección modal. Es la clave de estos modelos ya que, en base a las características de cada opción de viaje, la unidad de decisión (operador de mercancía/viajero) elige con una determinada probabilidad entre las diferentes opciones modales de transporte que existen.
4. Asignación a la red. De los flujos a cada tramo de la red modelizada, bien en número de vehículos, bien en toneladas (caso de mercancías). En este caso es necesario transformar toneladas en vehículos, teniendo en cuenta el ratio de toneladas/vehículo o utilizando unidades de medidas internacionales, como los contenedores estándar o TEUs. En cualquier caso, las unidades de medida han de ser representativas y estas aceptadas internacionalmente para evitar confusiones, ya sean físicas (longitud o peso), o monetarias. De esta forma, se obtienen resultados en cada tramo de red, en cada unidad correspondiente (viajes, número de envíos, toneladas, etc.), para cada modo de transporte.

La Figura 3.4 resume la estructura básica de un modelo de cuatro etapas, incluyendo

la etapa previa de zonificación del ámbito de estudio y construcción de la red, así como los datos de partida y las bases de datos en las que se fundamenta.

3.3.2. Especificaciones del Modelo clásico de cuatro etapas

A continuación, se resumen las especificaciones principales para los modelos de generación-atracción, distribución y reparto modal. La cuarta etapa de asignación, como su nombre indica, está basada en la asignación de los flujos de viajes (u otras unidades consideradas) a cada tramo de red utilizando diferentes algoritmos que cumplan el principio de equilibrio de Wardrop en la red modelizada.

Modelos de Generación-Atracción.

1. Objetivo del modelo: cuantificar el número de viajes que se generan y se atraen en cada zona de transporte.
2. Especificación del modelo:

$$G_i = \vec{\beta} * \vec{X}_i$$

$$A_j = \vec{\theta} * \vec{Y}_j$$

tal que $\sum_i G_i = \sum_j A_j$

1. Método de estimación: Mínimos Cuadrados Ordinarios
2. Entradas necesarias para estimar el modelo:
3. Vectores de viajes generados/atraídos *observados* por motivo de viaje
4. Variables de generación
5. Variables de atracción
6. Resultado del modelo: Vectores de viajes generados/atraídos *estimados* por motivo de viaje

Modelos de Distribución.

1. Objetivo del modelo: cuantificar de dónde a dónde se realizan los desplazamientos
2. Especificación del modelo:

$$V_{ij} = \frac{\alpha * G_i^\beta * A_j^\gamma}{C * G_{ij}^n},$$

donde, V_{ij} son los viajes generados en i y atraídos en j , y CG_{ij} es el coste generalizado entre las zonas i y j

3. Método de estimación: Mínimos Cuadrados Generalizados
4. Entradas necesarias para estimar el modelo:
5. Matriz de viajes *observados* generados/atraídos por motivo de viaje
6. Vectores de viajes generados/atraídos *estimados* por motivo de viaje
7. Matriz de costes generalizados
8. Resultado del modelo: Matriz de viajes *estimados* generados/atraídos por motivo de viaje

Modelos de Reparto Modal.

1. Objetivo del modelo: Cuantificar los viajes realizados en cada modo, basado en los valores observados (encuestas, datos de partida)
2. Especificación del modelo: Probabilidad (X): $P_i = \frac{e^{U_{ij}(X)}}{\sum_i e^{U_{ij}}}$
3. Método de estimación: Máxima Verosimilitud
4. Entradas necesarias para estimar el modelo:
5. Matriz de viajes observados por modos, de generados-atraídos por motivo de viaje
6. Matriz de costes generalizados por motivos de viaje

7. Resultado del modelo: Matriz de porcentajes, de cada modo de transporte, de viajes generados/atraídos por motivo de viaje

Como se ha indicado anteriormente, las matrices OD se pueden obtener bien mediante la transformación de los vectores de generados-atraídos, bien mediante encuesta o análisis de datos históricos existentes de fuentes oficiales. Si los orígenes y destinos de los viajes están perfectamente identificados, los procesos basados en encuestas resultan adecuados para conocer la tipología de la mercancía transportada y los perfiles de los viajeros que hacen uso de la red, por modo de transporte, eliminando la incertidumbre asociada a información estimada o agregada, la cual siempre es menos fiable.

Considerando lo indicado anteriormente, la evaluación de una infraestructura, funcional y económicamente (desde el punto de vista de la inversión y la rentabilidad), depende del número de usuarios que transiten por ella en sus vehículos, tanto desde el punto de vista de viajeros como del de las mercancías, por tanto, la construcción de los modelos de transporte permite realizar dichas evaluaciones y relacionarlas.

La siguiente figura representa de manera esquemática como es una red de transporte multimodal, en la que es necesario tener en cuenta los trazados de cada uno de los medios de transporte, sus puntos de encuentro con las terminales de intercambio modal necesarias y las características de estas.

Las variables mostradas son las que definen cada tramo, y están definidas en la siguiente Tabla.

De manera general los nodos de la red serán los puntos de origen y destino, los puertos y los puntos de intercambio modal. Los tramos deberán tener en cuenta todos los trazados existentes de carreteras y autopistas, ferrocarril existente, tramos marítimos y fluviales en explotación y todas las alternativas que se quieran evaluar no solo de ferrocarril sino de carretera o fluviales-marítimas.

Para el desarrollo de los siguientes apartados se toma de base la formulación matemática descrita en Ortúzar, J.D., y Willumsen, L. G. (2011)) en relación a los modelos de transporte, complementada con otros textos de modelos de regresión, funciones gamma, etc.

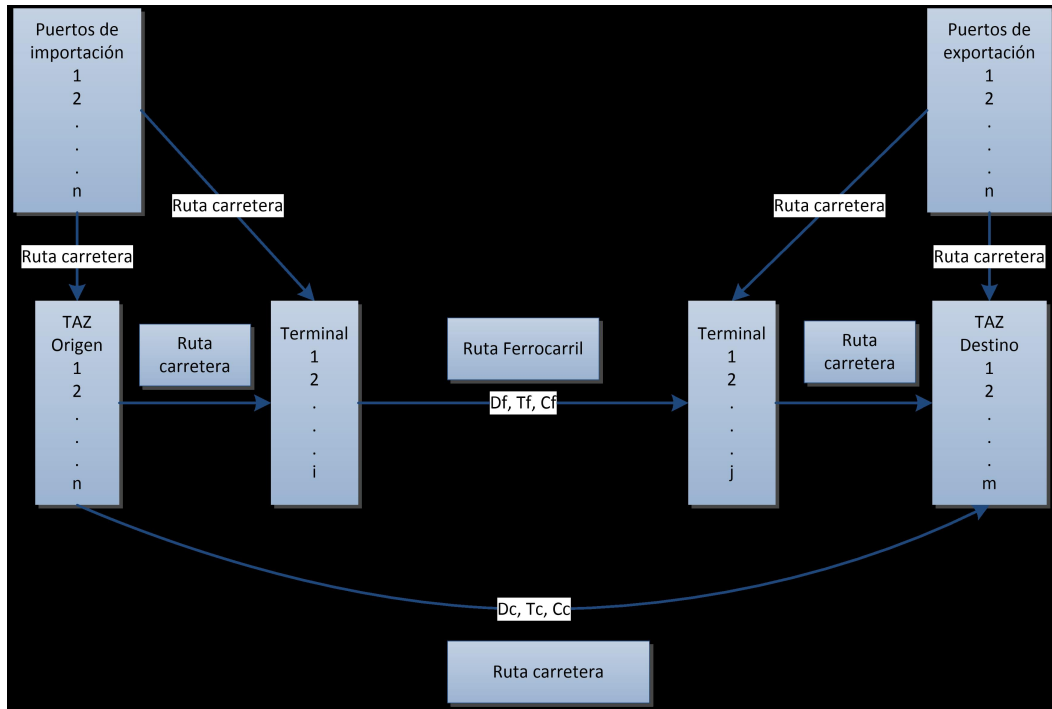


Figura 3.5: Red multimodal

Cuadro 3.1: Variables de las rutas

D_m	Distancia marítima
T_m	Tiempo marítimo
C_m	Coste marítimo
D_f	Distancia ferrocarril
T_f	Tiempo ferrocarril
C_f	Coste ferrocarril
D_c	Distancia carretera
T_c	Tiempo carretera
C_c	Coste carretera
T_{res}	Tiempo de espera
T_{op}	Tiempo operaciones
C_{op}	Coste operaciones

3.3.3. Etapa 1: Generación-Atracción de viajes.

Partimos de cuatro formulaciones matemáticas básicas: modelos de factores de crecimiento, modelos de regresión, modelos de análisis de tasas de viajes, y modelos de clasificación cruzada.

Formulación matemática: Recordando lo indicado anteriormente, los modelos de generación-atracción son tal que:

$$G_i = f(S_{i1}, S_{i2}, \dots, S_{in})$$

$$A_j = f(S_{j1}, S_{j2}, \dots, S_{jn})$$

Donde $f()$ son funciones asociadas a las variables socioeconómicas S_{ki} , como los modelos de regresión lineal múltiple indicados.

La condición que se debe cumplir en los modelos de generación atracción es la siguiente, siguiendo la condición de equilibrio de Wardrop de igualdad de flujos en un grafo:

$$\sum_i G_i = \sum_j A_j$$

Siendo G_i el número de viajes generados en la zona i , y A_j los viajes atraídos por la zona j . Para que esta condición se cumpla, normalmente es necesario usar un factor de corrección:

$$F = \frac{\sum_i G_i}{\sum_j A_j}$$

Por tanto, los viajes atraídos finales serán el FA_i .

Modelos de Factores de Expansión. Se desarrollaron al principio y ahora solo se utilizan para predicciones a corto plazo o para calcular los viajes externos a una zona de estudio.

El número de viajes futuros de una zona es el resultado de los viajes actuales

multiplicado por un factor de expansión, tal que:

$$T_i = F_i t_i,$$

donde,

1. T_i , viajes de la zona i
2. F_i , factor de expansión
3. t_i , viajes actuales de la zona i

Calcular el valor del factor de expansión es el objetivo primero, tal que F_i puede ser de la siguiente forma:

$$F_i = \frac{X_i^n Y_i^n Z_i^n}{X_i^0 Y_i^0 Z_i^0}$$

Siendo X_i , Y_i , Z_i variables socioeconómicas de la zona i (población, renta, PIB, . . .). El superíndice 0 corresponde al año base y el superíndice n al año futuro.

El procedimiento de estimación que propone este modelo es:

1. Determinar los valores actuales y futuros de las variables para cada zona y el volumen actual de viajes.
2. Estimar el valor del factor de expansión F_i .
3. Aplicar la formulación del modelo para cada zona en estudio.

Modelos de Regresión. Normalmente son multivariantes y de relación lineal o no lineal, en base a las variables seleccionadas.

El modelo de regresión lineal múltiple se muestra a continuación:

$$Y = k + b_1 X_1 + b_2 X_2 + \dots + b_n X_n,$$

con Y la variable dependiente (viajes producidos o atraídos por una zona); X_1, \dots, X_n las variables independientes; b_1, \dots, b_n los coeficientes de la regresión lineal; k el término de la variable dependientes no explicado por las variables independientes.

Aplicado a los modelos de transporte, en los modelos de generación la variable dependiente es el total de viajes generados por la zona, pudiendo ser a su vez, modelos de viajes zonales; de viajes basados en el hogar o desagregados (cuando al menos uno de los dos extremos del viaje es el hogar); o no basados en el hogar, en el caso complementario al anterior.

Las variables independientes de un modelo zonal son aquellas socioeconómicas de la zona en su conjunto; las utilizadas en un modelo desagregado son atienden a las variables de las propias viviendas.

La calibración del primero utiliza los datos existentes de la zonificación realizada, tal que cada observación es de una zona; mientras que en la calibración del modelo desagregado cada dato corresponde a cada vivienda de una muestra aleatoria tomada en el área en estudio.

Modelos de Regresión Lineal Múltiple. Sea un conjunto de n observaciones k -dimensionales, es decir,

$$x_1 = (x_{11}, \dots, x_{1k}), \dots, x_n = (x_{n1}, \dots, x_{nk})$$

Entonces, a partir de dos variables aleatorias (X, Y) , tal que X variable k dimensional, entonces se tiene un modelo lineal general si existen β_0 y β_j con $j=1, \dots, k$, verificando la siguiente condición:

$$y_i = \beta_0 + \sum_{j=1}^k x_{ij}\beta_j + \varepsilon_i, \quad i = 1, \dots, n$$

Con y_i la variable respuesta o dependiente; x_j las variables independientes o regresoras; β_j los coeficientes de regresión; y finalmente ε_i variables aleatorias incorreladas de media 0 y varianza σ^2 , para todo i (término de error).

De forma matricial el modelo lineal general quedaría:

$$Y = (Y_1, \dots, Y_n)^t$$

$$\beta = (\beta_0, \beta_1, \dots, \beta_k)^t$$

$$\varepsilon = (\varepsilon_1, \dots, \varepsilon_n)^t$$

$$X = \begin{pmatrix} 1 & x_{11} & & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & & & \vdots & \\ 1 & x_{n1} & & x_{n2} & \cdots & x_{nk} \end{pmatrix}_{n \times (k+1)}$$

Y entonces tendremos la siguiente forma matricial:

$$Y = X\beta + \varepsilon,$$

con $E(\varepsilon) = 0$ (vector de n ceros) y $V(\varepsilon) = \sigma^2 I_n$.

A partir de la siguiente expresión podemos conseguir el estimador de β , tal que para cada i se tienen las distancias verticales al cuadrado entre las verdaderas observaciones y_i y las estimadas a través del plano $\beta_0 + \sum_{j=1}^k x_{ij}\beta_j$ cuyo objetivo es minimizar $F(\beta)$ en β :

$$F(\beta) = \sum_{i=1}^n \left(y_i - \left(\beta_0 + \sum_{j=1}^k x_{ij}\beta_j \right) \right)^2 = (Y - X\beta)^t(Y - X\beta)$$

Este método corresponde al denominado de Mínimos Cuadrados, siendo:

$$\min_{\beta} F(\beta),$$

Se calcula:

$$\frac{\partial F}{\partial \beta} = -2X^tY + 2X^tX\beta = 0,$$

y se obtiene $X^tX\beta = X^tY$, de donde:

$$\hat{\beta} = (X^tX)^{-1}X^tY$$

Por lo tanto,

$$\hat{Y} = X\hat{\beta} = X(X^tX)^{-1}X^tY = HY,$$

siendo H matriz simétrica e idempotente.

Finalmente, se obtiene $\hat{\varepsilon}_i = Y_i - \hat{Y}_i$, para cada $i = 1, \dots, n$.

$$\hat{\varepsilon} = (\hat{\varepsilon}_1, \dots, \hat{\varepsilon}_n)^t = Y - \hat{Y} = Y - X\hat{\beta} = Y - X(X^tX)^{-1}X^tY = (I - H)Y.$$

3.3.4. Etapa 2: Distribución de viajes

El objetivo de los modelos de distribución de viajes es doble. Por un lado, ajustar en base a factores los viajes generados y atraídos de la etapa anterior, y por otro lado, transformar los viajes generados-atraídos en viajes origen-destino. Por tanto, al final de esta etapa podremos conocer los viajes entre las diferentes zonas de mi ámbito de estudio.

Los dos métodos más habituales en planificación de transporte son: 1) factores de crecimiento, basado en escalar una matriz existente aplicando factores multiplicativos derivados de estimaciones de generación-atracción asociadas a las celdas de la matriz; 2) modelo gravitacional, considerando la hipótesis de la existencia de medidas de impedancia entre las zonas como son: la distancia de viaje, el tiempo de viaje y su coste, así como otras funciones potenciales de impedancia que pueden ser usadas para derivar la atracción relativa de cada zona.

Así, en los modelos gravitacionales los viajes entre las zonas i y j pueden modelizarse en función de los viajes originados en las zonas i y el nivel de atracción y accesibilidad de la zona j con respecto a las demás zonas, considerando estas funciones de impedancia entre las zonas, las cuales están asociadas a la desutilidad del viaje. Este concepto de utilidad y desutilidad es de suma importancia en la etapa siguiente de reparto modal.

Siguiendo el símil de la ley de gravitación entre dos objetos, en este caso las zonas del ámbito, la “fuerza gravitatoria” (los viajes) decrece en función de la distancia que separa las zonas.

Ya que el resultante de la etapa anterior es un vector de viajes generados y un

Tabla 5.1. Estructura general de una matriz de viajes bi-dimensionales

Orígenes	Destinos					$\sum_j T_{ij}$
	1	2	3	...j	...z	
1	T_{11}	T_{12}	T_{13}	... T_{1j}	... T_{1z}	O_1
2	T_{21}	T_{22}	T_{23}	... T_{2j}	... T_{2z}	O_2
3	T_{31}	T_{32}	T_{33}	... T_{3j}	... T_{3z}	O_3
⋮						
i	T_{i1}	T_{i2}	T_{i3}	... T_{ij}	... T_{iz}	O_i
⋮						
z	T_{z1}	T_{z2}	T_{z3}	... T_{zj}	... T_{z2}	O_z
$\sum_i T_{ij}$	D_1	D_2	D_3	D_j	... D_z	$\sum_i T_{ij} = T$

Figura 3.6: Matriz de viajes. Fuente: Modelling Transport Manual (Ortúzar, J.D., y Willumsen, L. G. (2011))

vector de viajes atraídos, para cada zona, es necesario transformar estos vectores en forma matricial para una mejor interpretación.

A continuación se indican los elementos básicos del modelo:

1. **Matriz de viajes**, con orígenes en las filas y destinos en las columnas para representar el número de viajes que se producen de un origen i a un destino j . La notación es la siguiente:

- a) T_{ij} : número de viajes entre el origen i y el destino j
- b) O_i : número total de viajes con origen en la zona i
- c) D_j : número total de viajes atraídos por la zona j
- d) las celdas de cada una de las filas i contienen los viajes con origen en la zona i y destino en las zonas j .
- e) la diagonal principal corresponde a los viajes intrazonales.

Estas matrices de partida pueden filtrarse según el modo de transporte, la segmentación de la población, basados o no basados en el hogar, la franja horaria (AM, PM, Valle), el motivo del viaje, etc.

La suma de los viajes en una fila es igual al número total de viajes generados en esa zona; la suma de los viajes en una columna es el número total de viajes

atraídos a esa zona. Entonces:

$$\sum_i T_{ij} = O_i$$

$$\sum_j T_{ij} = D_j$$

Dependiendo de la cantidad y calidad de información de partida el modelo es doblemente restringido, si se cumplen ambas condiciones; o únicamente restringido, según se tenga sólo los generados o sólo los atraídos.

2. **Impedancias**, en términos de distancia, tiempo o unidades monetarias. En este punto se define el conocido como coste generalizado de viaje, que es una medida combinada de todos los principales atributos relacionados con la desutilidad de un viaje. Normalmente es una función lineal de los atributos del viaje aplicando un peso asociado a los coeficientes que representan su importancia relativa como la percibe el viajero. Así para el modo de transporte k se tiene:

$$c_{ij} = a_1 t_{ij}^v + a_2 t_{ij}^w + a_3 t_{ij}^t + a_4 t_{nij} + a_5 F_{ij} + a_6 \phi_j + \delta,$$

donde, si utilizamos variables como el tiempo, la tarifa (coste) y alguna posible penalización (por restricciones de viaje), entonces se definen:

- a) t_{ij}^v : tiempo de viaje en el vehículo entre i y j
- b) t_{ij}^w : tiempo andado hasta y desde las paradas o estaciones
- c) t_{ij}^t : tiempo de espera en las paradas
- d) t_{nij} : tiempo de intercambio
- e) F_{ij} : tarifa de ir de i a j
- f) ϕ_j : coste final (como podría ser, por ejemplo, el coste de un aparcamiento)
- g) δ : penalización modal (representa todos los demás atributos no incluidos en la medida generalizada, como la seguridad y la comodidad)
- h) $a_1, a_2, a_3, a_4, a_5, a_6$: pesos atribuidos a cada elemento del coste

Para poder hallar la matriz existen tres métodos: 1) el método del factor de crecimiento; 2) el modelo de distribución gravitacional; 3) la maximización de la entropía.

Modelo del Factor de Crecimiento. Se aplica si no se tiene información disponible relativa a las impedancias interzonales de la red: distancias, tiempos de viaje o costes generalizados.

Existen los siguientes tipos:

1. **Factor de crecimiento promedio:** distribuye los viajes futuros aplicando a la distribución actual el promedio de crecimiento esperado de la generación de viajes de una zona i y del crecimiento esperado de la atracción de viajes de otra zona j .

En el caso de que la única información disponible sea un grado general de crecimiento τ para el ámbito de estudio, se aplicará a cada celda de la matriz el siguiente valor:

$$T_{ij} = \tau \cdot t_{ij} \text{ para cada } i, j.$$

2. **Factor de crecimiento únicamente restringidos:** 1) si existe información del crecimiento esperado en viajes que se originan en cada zona se aplica este factor de crecimiento específico de origen (τ_i) a las correspondientes filas de la matriz de viajes; 2) si existe información del crecimiento esperado de viajes atraídos en cada zona se aplica el factor de crecimiento específico de destino (τ_j) a las correspondientes columnas, tal que:

$$T_{ij} = \tau_i t_{ij} \text{ para los factores específicos de origen}$$

$$T_{ij} = \tau_j t_{ij} \text{ para los factores específicos de destino}$$

3. **Factor de crecimiento doblemente restringidos:** en este caso se conoce el número de viajes futuros generados y atraídos en cada zona.

Estos métodos requieren el cálculo de un conjunto de coeficientes de corrección intermedios que se aplican a cada celda de entrada, en cada fila o columna

según convenga. En el caso de asignar estas correcciones a cada fila, se calculan los totales para cada columna y se comparan con los valores objetivos. Si las diferencias son significativas, se calculan y aplican nuevos coeficientes de corrección.

Entonces se aplica el siguiente proceso iterativo:

- a) Considerar a todos los $b_j = 1$ y resolver para los a_i ; es decir, hallar los factores de corrección a_i que cumplan las restricciones relativas a la generación de viajes.
- b) A partir de los últimos valores obtenidos para los a_i , hallar los b_j ; de manera que se cumplan las restricciones relativas a la atracción de viajes.
- c) Manteniendo fijos los valores de los b_j , resolver para los a_i y repetir los pasos 2 y 3 hasta que los cambios sean suficientemente pequeños.

Modelo de distribución gravitacional. Estima los viajes para cada celda en la matriz sin utilizar directamente el esquema de viajes observados; por tanto, se les llama algunas veces como métodos sintéticos como oposición a los modelos de factor de crecimiento.

Su primera formulación y más simple es tal que: $T_{ij} = \frac{\alpha P_i P_j}{d_{ij}}$

En este caso se consideran como variables aleatorias de referencia P_i y P_j , las poblaciones de las zonas del ámbito de estudio; d_{ij} la distancia entre i y j ; y α un factor de proporcionalidad.

Posteriormente se incluyeron mejoras como la utilización de los totales de los viajes finales (O_i y D_j), en lugar de las poblaciones totales, y un parámetro n para la calibración de la potencia de d_{ij} . Este parámetro n , no se restringió a los números enteros y ciertos estudios estimaron su valor entre 0.6 y 3.5.

En este caso el modelo resultante es:

$$T_{ij} = \alpha O_i D_j f(c_{ij}),$$

donde $f(c_{ij})$ es una función de costes generalizados de viaje con uno o más parámetros para la calibración; esta función se conoce como *función de disuasión* que re-

presenta un grado de penalización al viaje al aumentar el coste o la distancia. Las formas más utilizadas de esta función son:

1. Función exponencial: $f(c_{ij}) = \exp(-\beta c_{ij})$
2. Función potencial: $f(c_{ij}) = c_{ij}^{-n}$
3. Función combinada: $f(c_{ij}) = c_{ij}^{-n} \exp(-\beta c_{ij})$

De cara a asegurar el cumplimiento de las restricciones de los totales por orígenes y destinos, se reemplaza el factor α de proporcionalidad por dos conjuntos de factores de balanceado: A_i y B_j obteniendo entonces el Modelo Gravitacional Doblemente restringido en su forma habitual, tal que:

$$T_{ij} = A_i O_i B_j D_j f(c_{ij}).$$

El *modelo únicamente restringido* por origen o por destino, considera uno de los conjuntos de factores de balanceado A_i o B_j , igual a 1.

1. Para un modelo restringido por orígenes, $B_j = 1$ para todo j , y $A_i = \frac{1}{\sum_j D_j f(c_{ij})}$
2. Para el caso restringido por destinos sería análogo cambiando las posiciones de A_i y B_j .

El *modelo doblemente restringido*, considera los valores de los factores de balanceado:

$$A_i = \frac{1}{\sum_j B_j D_j f(c_{ij})}$$

$$B_j = \frac{1}{\sum_i A_i O_i f(c_{ij})}$$

Entonces los factores de balanceo son interdependientes, es decir, para calcular un conjunto de factores se necesita utilizar los valores del otro conjunto y viceversa. Por ese motivo es necesario realizar iteraciones análogas a las de **Furness** que, en la práctica, producen buenos resultados.

Luego, dado un conjunto de valores para la función de fricción (o impedancia o de resistencia al viaje) $f(c_{ij})$, el proceso inicializa con todos los $B_j = 1$, calculando los valores de A_i mediante la primera fórmula descrita anteriormente. A continuación, con estos valores hay que reestimar de nuevo los B_j y repetir estos pasos iterativamente hasta que el proceso converja.

Una versión más general de la función de fricción puede incluir valores empíricos que dependen sólo del coste generalizado de viaje. Los costes de viaje se agregan entonces en un número pequeño de intervalos de coste (10 ó 15), indicados con el superíndice m , de tal forma que la función de fricción queda así:

$$f(C_{ij}) = \sum_m F^m \delta_{ij}^m$$

Donde F^m es el valor medio del coste para el intervalo m , mientras que δ_{ij}^m es igual a 1 si el coste de viaje entre i y j pertenece al intervalo m , e igual a 0 en otro caso.

La formulación de la función potencial y exponencial tienen solamente un parámetro a calibrar.

La formulación de la combinada tiene dos parámetros.

Mientras que la fórmula anteriormente indicada tiene tantos parámetros como intervalos. Estos parámetros se estimados tal que los resultados obtenidos reproduzcan lo más fielmente posible la distribución de longitudes (costes) de los viajes (DLV o, en inglés, Trip Length Distribution, TLD) observados. El programa de modelización de transporte TransCAD incluye esta opción.

De tal forma que se calcula la matriz O-D a través de la siguiente fórmula:

$$T_{ij} = k G_i A_j f(C_{ij}) = \frac{G_i A_j f(C_{ij})}{\sum A_j f(C_{ij})},$$

con:

1. T_{ij} : viajes con origen en la zona i y destino en la zona j .
2. G_i : total de viajes generados en la zona i
3. A_j : total de viajes atraídos por la zona j

4. $f(C_{ij})$: función de coste entre i y j
5. k : constante de proporcionalidad

En este caso se asume que $B_j = 1$ y la función de coste es el tiempo (minutos) de viaje entre la zona i y la zona j .

Usualmente se puede usar la notación V_{ij} en lugar de T_{ij} .

Maximización de la entropía. Se define la entropía como la probabilidad relacionada con la incertidumbre respecto a la información de que se dispone inicialmente. Se consideran tres niveles de información:

1. Macro-estado: Viajes generados y atraídos por cada zona. O_i y D_j .
2. Meso-estado: Viajes realizados entre una zona i y una zona j .
3. Micro-estado: Viajes de un individuo desde un origen i a un destino j .

El número de micro-estados $W\{T_{ij}\}$ asociados al meso-estado T_{ij} viene dado por:

$$W\{T_{ij}\} = \frac{T!}{\pi_{ij} T_{ij}!}$$

Suponiendo a todos los micro-estados igualmente probables, el meso-estado más probable es aquel que puede generarse por un gran número de formas.

Entonces, se necesita una técnica para identificar los valores que maximizan W en la expresión anterior. Se considera entonces tomar logaritmos tal que:

$$\log W = \log \frac{T!}{\pi_{ij} T_{ij}!} = \log T! - \sum_{ij} \log T_{ij}!$$

Utilizando la aproximación de Stirling, entonces tengo $\log X! = X \log X - X$

$$\log W = \log T! - \sum_{ij} (T_{ij} \log T_{ij} - T_{ij})$$

Y como $\log T!$ es una constante, el problema se reduce a maximizar la expresión

$$\log W' = - \sum_{ij} (T_{ij} \log T_{ij} - T_{ij}),$$

que es la que se conoce como **función de entropía**.

3.3.5. Etapa 3: Reparto o elección modal

En la anterior (segunda) etapa se obtiene como resultado final una matriz OD con los viajes entre cada una de las zonas. El objetivo de esta tercera etapa es cuantificar cuántos de esos desplazamientos se producen en cada modo de transporte, calculando las probabilidades de uso de cada uno de ellos.

El motivo es que conocer la elección del modo de transporte en escenarios futuros es clave en la planificación del transporte para la toma de decisiones, dado que condiciona totalmente las inversiones en infraestructuras a largo plazo.

Además, conocer la elección modal influye en la eficiencia general del sistema del transporte en el escenario actual, en la cantidad del espacio urbano dedicado a las funciones del transporte, así como en el conjunto de alternativas disponibles o no para los viajeros. Todo esto está relacionado con las políticas en materia de movilidad, sociales, urbanísticas y de medio ambiente.

Los modelos de reparto modal se utilizan para analizar y predecir las elecciones que realizan los individuos o grupos de individuos en cuanto al modo de transporte del viaje. Normalmente, el objetivo es predecir la cuota porcentual de cada modo en todas las relaciones origen-destino con el fin de obtener una matriz O-D para cada modo.

Esto implica que es importante desarrollar y utilizar modelos que sean sensibles a aquellos atributos del viaje que influyen en las elecciones individuales del modo. Los modelos aplicados después de la fase de distribución permiten tener en cuenta las características del viaje y de las alternativas disponibles para realizarlo, si bien es difícil incluir las características de los viajeros que ya se han agregado en la matriz de viajes de la etapa anterior.

Maximización de la utilidad. Los factores que influyen en la elección modal se pueden clasificar en tres grupos:

1. **Personal:** disponibilidad de vehículo, estructura del hogar, ingresos, ...

2. **Viaje.** La elección del modo está fuertemente influenciada por el propósito del viaje y por el período del día en el que se realiza el viaje. Así, los viajes realizados en las últimas horas del día se realizan con mayor dificultad en transporte público.
3. **Modo de transporte,** que se dividen en dos categorías: 1) formada por factores cuantitativos como el tiempo relativo del viaje: tiempo de viaje a bordo del vehículo, de espera y a pie por cada modo o por el coste del viaje; 2) formada por factores cualitativos, menos fáciles de medir, como la confiabilidad y regularidad.

Los modelos de elección modal pueden ser agregados (si se basan en datos a nivel zonal e interzonal) o desagregados (si se basan en datos familiares o individuales). Los modelos agregados generalmente utilizan el concepto de coste generalizado mientras que los modelos desagregados pueden incluir potencialmente la mayoría de los factores anteriormente mencionados.

Los modelos de elección discreta afirman: *“La probabilidad de que los individuos elijan una determinada alternativa es función de sus características socioeconómicas y de la relativa atractividad de la alternativa.”*

Para representar la atractividad de la alternativa se utiliza el concepto de utilidad. La utilidad medible u observable se define generalmente como una combinación lineal de variables donde cada variable representa un atributo de la alternativa o del viajero, en tanto que los coeficientes representan la influencia relativa de cada atributo, es decir, la contribución que cada variable aporta al agrado total producido por cada alternativa.

La constante de la función de utilidad se puede interpretar como la representación de la influencia neta de todas las características, tanto del individuo como de la alternativa de transporte, no observadas o no explícitamente incluidas en dicha función de utilidad. La constante específica puede incluir, por ejemplo, elementos tales como el confort o la fiabilidad, que son variables nada fáciles de medir o de observar.

De acuerdo con el modelo, para poder predecir si una alternativa es elegida, el valor de su utilidad se ha de comparar con el valor de las utilidades de las opciones alternativas y transformarse en un valor de probabilidad entre 0 y 1. Para ello, existe

una gran variedad de transformaciones matemáticas, cuyas gráficas tienen forma de S y de entre las cuales se destacan las siguientes:

1. Logit:

$$P_1 = \frac{\exp(V_1)}{\exp(V_1) + \exp(V_2)}.$$

2. Probit:

$$P_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{V_1 - V_2 + x_1} \frac{\exp\left\{-\frac{1}{2(1-p^2)} \left[\left(\frac{x_1}{\sigma_1}\right)^2 - \frac{2px_1x_2}{\sigma_1\sigma_2} + \left(\frac{x_2}{\sigma_2}\right)^2\right]\right\}}{2\pi\sigma_1\sigma_2\sqrt{(1-p^2)}} dx_2 dx_1,$$

donde la matriz de covarianzas de la distribución Normal asociada a este último modelo tiene la forma siguiente:

$$\Sigma = \begin{pmatrix} \sigma_1^2 & p\sigma_1\sigma_2 \\ p\sigma_1\sigma_2 & \sigma_2^2 \end{pmatrix}.$$

Los modelos de elección discreta no se pueden calibrar utilizando técnicas clásicas de ajuste de curvas, como, por ejemplo, el método de los Mínimos Cuadrados, porque su variable dependiente P_i es una probabilidad no-observada (entre 0 y 1), mientras que las observaciones son las elecciones realizadas por los individuos (que son solamente 0 ó 1); las únicas excepciones al respecto son los modelos para grupos homogéneos de individuos, o cuando el comportamiento de cada individuo se registra en varias ocasiones, porque, de hecho, las frecuencias observadas de elección son también variables comprendidas entre 0 y 1.

Spear (1977) realizó una conveniente síntesis de algunas propiedades importantes de estos modelos:

Por su parte, los modelos desagregados de demanda (DM) se basan en teorías de comportamiento individual y no constituyen analogías físicas de ningún tipo. Por ese motivo, cuando intentan explicar dicho comportamiento, presentan una importante ventaja potencial respecto a los modelos convencionales, en el sentido de que es más probable que sean estables en el tiempo y en el espacio.

Los DM se estiman utilizando datos individuales, lo que implica que:

1. En lo referente a la utilización de información pueden ser más eficientes que los modelos convencionales; de hecho, requieren un número menor de datos ya que la elección de cada individuo puede ser utilizada como una observación.
2. Asimismo, el utilizar datos individuales hace que se pueda tener en cuenta toda la variabilidad inherente a dichas informaciones.
3. Una menor probabilidad de que se vean afectados por distorsiones debidas a la correlación entre unidades agregadas.

Los modelos desagregados son probabilísticos; ello implica que como proporcionan la probabilidad de elegir cada alternativa, sin indicar cuál se selecciona, se deben utilizar conceptos básicos de la teoría de las probabilidades, tales como:

1. El número esperado de personas que utilizan una cierta alternativa de viaje es igual a la suma, sobre todos los individuos, de la probabilidad de elección de dicha alternativa:

$$N_i = \sum P_{in_n}$$

2. Es posible modelizar un conjunto de decisiones independientes considerando a cada una como una elección condicionada; en este caso separadamente, las probabilidades resultantes pueden ser multiplicadas para obtener las probabilidades conjuntas, tal que:

$$P(f, d, m, r) = P(f) P(d/f) P(m/d, f) P(r/m, d, f)$$

Las variables explicativas incluidas en el modelo pueden tener coeficientes explícitamente estimados. Contrariamente a lo que sucede en el caso del coste generalizado de los modelos convencionales, en los que la función está generalmente limitada y presenta numerosos parámetros fijos, en la función de utilidad es posible, en principio, insertar un número cualquiera de variables explicativas con cualquier tipo de especificación.

Al modelizar el comportamiento de los distintos usuarios del sistema sería necesario conocer todas las características que han tenido en cuenta a la hora de hacer su elección. Esto es imposible, ya que no se llega a tener nunca una información completa acerca de lo que los individuos consideran a la hora de elegir una alternativa. Por esta razón la utilidad de las alternativas U_{iq} se representa con la siguiente ecuación:

$$U_{iq} = V_{iq} + \varepsilon_{iq}.$$

V_{iq} : Parte medible de la Utilidad, Utilidad sistemática o representativa, que es función de los atributos medidos X , siendo:

$$V_{iq} = \sum \theta_{ki} X_{kiq}$$

donde θ se considera constante para todos los individuos.

ε_{iq} : Parte aleatoria que refleja la idiosincrasia y gustos particulares de cada individuo, además de errores de medición y observación cometidos por parte del equipo investigador.

Es importante enfatizar la existencia de dos puntos de vista en la formulación de este problema: 1) el individuo que sopesa todos los elementos de interés (sin aleatoriedad) y a continuación elige la alternativa más conveniente; 2) el modelizador, el cual observando solamente algunos de los elementos, necesita incluir los residuos ε para explicar lo que de otra manera podría constituir un comportamiento no racional.

Por tanto, el individuo elige la alternativa que le proporciona su máxima utilidad, es decir elige A_j si, y sólo si:

$$U_{jp} \geq U_{iq} \quad \forall A_i \in A(q)$$

Es decir,

$$V_{jq} - V_{iq} \geq \varepsilon_{iq} - \varepsilon_{jq}.$$

Maximización de la entropía. El enfoque de maximización de la entropía puede utilizarse para generar modelos simultáneos de distribución y elección modal. Para

ello, es necesario formular el problema de maximización de la entropía; es decir maximizar la siguiente expresión:

$$\log W(T_{ij}^k) = - \sum_{ijk} (T_{ij}^k \log T_{ij}^k - T_{ij}^k),$$

sujeto a

$$\sum_{jk} T_{ij}^k - O_i = 0$$

$$\sum_{ik} T_{ij}^k - D_j = 0$$

$$\sum_{ijk} T_{ij}^k C_{ij}^k - C = 0$$

Lo cual nos lleva a la siguiente solución:

$$T_{ij}^k = A_i O_i B_j D_j \exp(-\beta C_{ij}^k)$$

$$P_{ij}^l = \frac{T_{ij}^l}{T_{ij}} = \frac{\exp(-\beta C_{ij}^l)}{\sum_k \exp(-\beta C_{ij}^k)},$$

donde P_{ij}^l es la proporción de viajes realizados entre i y j con el modo l (k son los diferentes modos).

Las principales propiedades que presenta P_{ij}^l son las siguientes:

1. Genera una curva en forma de “S” conocida como curva de partición modal.
2. Si $C_1 = C_2 = \dots = C_n$ para todos los valores de k entonces $P_1 = P_2 = \dots = P_n = \frac{1}{n}$.

Es importante mencionar que en esta formulación β juega un doble papel, por un lado, como parámetro que controla la dispersión en la elección modal y por otro en la elección de diferentes destinos que se encuentran a distancias diferentes del origen. Por este motivo, a veces se prefiere no tener un solo parámetro sino tener varios, apareciendo así un modelo más práctico de la forma:

$$T_{ij}^{kn} = A_i^n O_i^n B_j D_j \exp(-\beta_n K_{ij}^n) \frac{\exp(-\lambda_n C_{ij}^k)}{\sum_k \exp(-\lambda_n C_{ij}^k)}$$

Donde K_{ij}^n es el coste compuesto de viaje entre i y j percibido por el individuo de tipo n . El coste compuesto puede ser especificado de muchas formas diferentes pero la mejor forma suele ser tomar su promedio ponderado:

$$K = \sum_k P^k C^k.$$

También se ha establecido que la única especificación correcta compatible con la teoría predominante del comportamiento de elección racional es:

$$K_{ij}^n = \frac{-1}{\lambda_n} \log \sum_k \exp(-\lambda_n C_{ij}^k)$$

Donde se tiene que satisfacer la restricción:

$$\beta_n \leq \lambda_n$$

Y, además, esta medida de coste compuesto tiene las siguientes propiedades:

1. $K = \min_k (C^k)$
2. $\lim_{\lambda \rightarrow \infty} K = \min_k (C^k)$
3. $\frac{dK}{dC^k} = P^k$

3.3.6. Etapa 4: Asignación a la red

En esta fase se utiliza la información de la matriz OD de viajes de la anterior etapa para asignarlos a la red modelizada del ámbito de estudio. Es decir, se obtiene como resultado final del modelo la cantidad de viajes que pasan por cada uno de los arcos en diferentes modos. El objetivo es encontrar las rutas óptimas mediante diferentes algoritmos que dependen del tipo de transporte utilizado.

1. **Transporte privado.** Se sigue el principio de equilibrio de Wardrop. Bajo condiciones de equilibrio el tráfico se distribuye en la red tal que todas las rutas usadas entre un par origen-destino tienen el mismo coste y cualquier otra ruta tiene mayor coste.
2. **Transporte público (TP).** Se aplica un procedimiento probabilístico que reparte los flujos en TP entre las rutas mono o multimodales lógicas que puedan generarse para cada par origen-destino.

La notación utilizada es la siguiente,

T_{ijr} : Número de viajes entre i y j por el recorrido o ruta r .

V_a : Flujo en el arco a .

$C(V_a)$: Relación coste-flujos para el arco a .

$c(V_a)$: Coste efectivo para un nivel específico de flujo V_a . En el caso de que $V_a = 0$ se trata del coste a flujo libre.

c_{ijr} : Coste de viajar de i a j por el recorrido r .

$$\delta_{ijr}^a = \begin{cases} 1 & \text{si el arco } a \text{ pertenece al recorrido } r \text{ desde } i \text{ a } j \\ 0 & \text{en otro caso} \end{cases}$$

n : Indica una iteración particular en los métodos iterativos.

$*$: Indica el valor óptimo

Durante esta etapa se utilizan una serie de principios para cargar una matriz de viajes a la red y obtener unos flujos en los arcos. Los objetivos son los siguientes:

1. Obtener resultados agregados de red.
2. Estimar costes de viaje entre zonas para un nivel de demanda dado.
3. Conseguir valores razonables de flujos en los arcos e identificar los arcos que estén más congestionados.

4. Estimar los recorridos utilizados para cada par O-D.
5. Analizar qué pares O-D utilizan arcos o rutas particulares.
6. Obtener los flujos de vehículos que realizan giros para poder así diseñar futuras intersecciones.

Los requerimientos básicos para esta etapa son:

1. Una matriz de viajes que exprese la demanda estimada. Es común que esta matriz se refiera a una franja horaria punta en un área congestionada. En el caso de redes no congestionadas se usan matrices diarias.
2. Una red: los arcos y sus propiedades.
3. Principios y reglas de selección de los recorridos relevantes para el problema.

Se atribuye a dos factores que los individuos elijan diferentes recorridos cuando se desplazan entre el mismo par O-D: las diferencias individuales de los individuos sobre cuál es el mejor recorrido; los efectos que tiene la congestión, que hace que los costes de los trayectos más cortos aumenten.

La construcción de un buen árbol de asignación en esta etapa en los métodos de asignación sirve para ahorrar tiempo y costes. En los algoritmos que veremos en el siguiente apartado, indicaremos con $d_{A,B}$ como la longitud (esto es, el coste) de un arco entre A y B en la red. La ruta se define por una serie de nodos unidos. Con d_A se denota la distancia mínima para alcanzar el centroide A desde el nodo raíz S del árbol.

3.3.7. Métodos de asignación

Los métodos de asignación son los siguientes:

1. **Asignación todo o nada.** Este método supone que no hay congestión lo que se traduce en que los costes de los arcos son fijos. Además, se supone que todos los individuos tienen que percibir de la misma manera el coste. Esto es, todos los individuos que se desplazan de i a j , tienen que elegir la misma ruta.

Estas suposiciones son razonables en redes poco congestionadas y con pocas alternativas de recorridos.

Todos los algoritmos de carga de la red comienzan con una fase inicial, hay que determinar los flujos $V_{A,B}$ en los arcos entre los nodos A y B. Aquí se comienza con $V_{A,B} = 0$. Podemos aplicar dos métodos para resolver el problema:

a) Par por par: se parte de un origen y un destino a la vez. Inicialmente, $V_{A,B} = 0$, luego para cada par (i, j) :

Se fija B igual al destino j

Si (A, B) es el arco predecesor de B entonces $V_{A,B} = V_{A,B} + T_{ij}$

Se fija B igual al destino A

Si $A = i$ finalizamos. En otro caso, se vuelve al segundo paso.

b) En cascada: a partir del origen i se carga sobre los arcos los flujos acumulados de los nodos siguiente los árboles de mínimo coste. Sea V_A el flujo acumulado en A:

Se fijan todos los $V_A = 0$ excepto para el destino en que $V_j = T_{ij}$.

Se fija B igual a la mayor distancia al nodo más lejano desde i .

$$V_A = V_A + V_B$$

$$V_{A,B} = V_{A,B} + V_B$$

Se fija B igual al siguiente nodo más lejano. Si $B = i$ el origen ha sido alcanzado y se vuelve a comenzar por el origen siguiente. En otro caso, se vuelve al tercer paso.

2. **Métodos estocásticos.** Estos métodos ponen en manifiesto la variabilidad por parte de los individuos de la percepción de los costes y tratan de minimizar la distancia, el tiempo de viaje y los costes. Es necesario tener en cuenta el mejor recorrido, pero también los recorridos siguientes a este, hablando en términos de costes calculados por el modelizador. Se denominan los “second best routes” y el número de rutas alternativas entre cada par origen-destino es muy grande.

Se usan dos métodos importantes: 1) el primero de ellos serían los métodos que se basan en la simulación, utilizando en este caso métodos de Monte Carlo

para representar las diferentes percepciones de los usuarios sobre los costes; 2) el segundo se basa en proporciones, estos asignan flujos a las diferentes rutas alternativas de recorrido de proporciones calculadas utilizando expresiones tipo logit.

3. **Asignación con congestión.** En la asignación con congestión nos centraremos en modelos que buscan aproximarse a las condiciones de equilibrio de Wardrop. Con estos modelos se concentra la atención en las restricciones de capacidad. Los modelos con restricciones de capacidad usan funciones que relacionan el flujo con los costes de viaje sobre los arcos.

El equilibrio de Wardrop: “*Bajo condiciones de equilibrio, el tráfico se distribuye en las redes congestionadas de modo tal que ningún viajero puede reducir su propio coste de viaje cambiando el recorrido*”.

Se usa el indicador

$$\delta = \frac{\sum_{ijr} T_{ijr} (C_{ijr} - C_{ijr}^*)}{\sum_{ij} T_{ij} C_{ij}^*}$$

para valorar en qué medida la solución se acerca al equilibrio de Wardrop. Se trata de una medida sobre coste total debido al hecho de usar recorridos peores, esto es, diferentes a los del coste mínimo.

El segundo principio de Wardrop “*Bajo condiciones de equilibrio social en redes congestionadas, el tráfico debería distribuirse de tal modo que los costes medios (o totales) de viaje sean mínimos*”.

3.4. Consideraciones de los modelos de elección modal

Por un lado, un aspecto clave en el desarrollo de modelos de transporte es la estimación del reparto modal entre una serie de alternativas. Por otro lado, la decisión del modo de transporte a la que se enfrenta un individuo es de naturaleza discreta. Por lo tanto, los modelos de elección discretos son a menudo los elegidos

para el análisis del modo de decisión. Estos predicen las elecciones hechas por una unidad de decisión dentro de un conjunto discreto de alternativas. Sin embargo, los modelos de elección discretos están formulados como modelos estocásticos, en los que la probabilidad, con la que determinada respuesta es observada, es una función de un conjunto de variables explicativas.

Los modelos de elección discreta persiguen precisamente estimar cómo distintos decisores escogen entre una serie de opciones que pueden corresponder a distintos modos de transporte, combinaciones de modos o configuraciones de un modo de transporte en función de un conjunto de factores que determinan la elección (Ortúzar, J.D., y Willumsen, L. G. (2011)).

En el transporte de pasajeros y mercancías, los dos factores más determinantes en la elección modal son el tiempo y el coste de viaje por cada opción, aunque debido a la disparidad existente entre los criterios seguidos por usuarios distintos, se considera adecuado el uso de modelos probabilísticos frente a modelos determinísticos de forma que se contemple la variabilidad en las decisiones de los usuarios del medio de transporte. Por tanto, el modelo de elección ha de proporcionar la probabilidad de uso de una alternativa, y esa probabilidad es usada en un modelo de transporte (agregado usualmente), como la proporción de usuarios que toman la alternativa.

Con el objetivo de evaluar la captación de tráfico que proporciona un determinado modo de transporte frente a sus alternativas, se emplea una metodología basada en el empleo de modelos de elección discreta.

Los más populares aplicados en la práctica son los modelos logit, los cuales pueden subdividirse en varios tipos. Los modelos de elección discretos son en muchos aspectos un sustituto para los modelos de regresión cuando la variable dependiente es cualitativa o categórica y no continua, y dichos modelos de regresión no se ajustan a la modelización de variables dependientes discretas por violaciones de los supuestos de mínimos cuadrados (Aldrich and Nelson, (1984)). Los principales modelos logit utilizados son los Multinomial Logit (MNL), los Binary Logit (BL) y los modelos logit anidados NML (Nested Logit Models) que son variaciones que parten de hipótesis más complejas. Además de estos modelos, también podemos destacar la utilización de las Redes Neuronales (RN) o Neural Networks (NN), aunque en menor medida. En el caso de los MNL, se trata de una metodología robusta y ampliamente extendi-

da en la práctica, que permite determinar las probabilidades de que diversos agentes tomen unas decisiones u otras en función de un conjunto de factores determinantes de la elección.

Concretando algo más, los modelos de elección discreta están basados en la teoría de la utilidad aleatoria, cuyas asunciones se pueden resumir en (Ortúzar and Willumsen, (2011)): existe una población homogénea formada por un conjunto de individuos que disponen de información “perfecta”; el conjunto total de alternativas (de cada modo de transporte) está disponible para ellos; cada alternativa de transporte (j) tiene una utilidad de red por cada individuo (p), que se define como la suma de dos componentes, una función medible de ciertos factores que afectan a la elección del modo de transporte (V_{jp}), y una componente aleatoria de cada individuo tomada como una medida del error observado (ε_{jp}).

Por lo tanto, se tiene una expresión funcional de la forma:

$$U_{jp} = V_{jp} + \varepsilon_{jp}.$$

Según una de las principales referencias bibliográficas sobre modelización de transporte (Ortúzar, J.D., y Willumsen, L. G. (2011)), por un lado tenemos que: la elección del modo de transporte dentro de la red construida, representa el elemento más importante en la planificación de los transportes y en la toma de decisiones; tanto para modelos de pasajeros como de mercancías, es la etapa crucial de cara a las proyecciones a futuro; influye en la eficiencia general del sistema de transportes, en la cantidad de espacio urbano dedicado a las funciones del transporte, así como en el conjunto de alternativas disponibles;

Además, se tiene que la probabilidad de que los individuos elijan una determinada alternativa es función de sus características socioeconómicas y de la relativa atractividad de la alternativa. Para representar la atractividad de la alternativa se utiliza el concepto de utilidad (éste es un artificio teórico convenientemente definido en forma tautológica como lo que el individuo intenta maximizar). Las alternativas per se no producen utilidad, sino que la utilidad se deriva, Lancaster, (1966), de las características de las alternativas y de las características de los individuos. La utilidad medible u observable se define generalmente como una combinación lineal de variables. Hay que comparar el valor de las utilidades de cada alternativa

y transformarlos en un valor de probabilidad entre 0 y 1, utilizando habitualmente transformaciones matemáticas entre las que destacan los modelos Logit y Probit.

Logit:

$$P_1 = \frac{\exp(V_1)}{\exp(V_1) + \exp(V_2)}$$

Probit:

$$P_1 = \int_{-\infty}^{\infty} \int_{-\infty}^{V_1 - V_2 + x_1} \frac{\exp\left\{-\frac{1}{2(1-\rho^2)} \left[\left(\frac{x_1}{\sigma_1}\right)^2 - \frac{2\rho x_1 x_2}{\sigma_1 \sigma_2} + \left(\frac{x_2}{\sigma_2}\right)^2\right]\right\}}{2\pi\sigma_1\sigma_2\sqrt{(1-\rho^2)}} dx_2 dx_1$$

Siendo los más utilizados de estos los siguientes:

Multinomial Logit (MNL) (Ortúzar, J.D., y Willumsen, L. G. (2011)):

$$P_{iq} = \frac{\exp(\beta V_{iq})}{\sum_{A_j \in A(q)} \exp(\beta V_{jq})}$$

Hierarchical Logit (HL), o Modelo Logit Anidados (Williams, (1977); Daly y Zachary, (1978)):

$$P(d, m) = \frac{\exp\{\beta(V_d + V_d^*)\} \exp(\lambda V_{dm})}{\sum_{d'} \exp\{\beta(V_{d'} + V_{d'}^*)\} \sum_{m'} \exp\{\lambda V_{d'm'}\}}$$

con

$$V_d^* = \left(\frac{1}{\lambda}\right) \log \sum_{m^*} \exp(\lambda V_{dm^*})$$

En opinión de Ortúzar y Willumsen, (2011): “Toda esta teoría se basa en la hipótesis de que la unidad de decisión, el viajero IDEAL, es racional, egoísta, y sus gustos NUNCA cambian. Maximizando su utilidad mediante análisis cuidadosos y reflexivos.”; “Sin embargo, el viajero REAL, es parcialmente racional, pero también es emocional y colaborador. No puede usar TODAS las alternativas, por lo que usa reglas heurísticas para decidir: Le importan más los cambios que los valores absolutos; Tiene una sensibilidad decreciente a los cambios de utilidad (5 minutos

son importantes en un viaje de 30 pero no en uno de 3 horas); Es enemigo de las pérdidas; No reacciona inmediatamente”.

Dicho esto, dado el grado de incertidumbre y la información a priori en la toma de decisiones por parte de la unidad de decisión (en base a recuerdos y experiencias pasadas), por qué no plantear otro tipo de modelos para analizar la elección de alternativas?, más allá de modelos Logit Multinomiales (MNL) o Logit Jerárquicos (HL).

¿Por qué no optimizar la etapa de reparto modal utilizando Redes Bayesianas?, dados los buenos resultados de las mismas en los últimos años en diferentes trabajos de investigación (utilizando una gran cantidad de variables), y dada la relevancia de esta etapa en los modelos de transporte.

3.4.1. Multinomial Logit (MNL)

Este modelo formula la probabilidad de que una unidad de transporte elija una alternativa dada de un conjunto de posibles alternativas. Esta probabilidad viene dada por:

$$P_n(i) = \text{prob}(Y_n) = \frac{e^{V_{ni}}}{\sum_{j \in C_n} e^{V_{nj}}},$$

donde:

$P_n(i)$ es la probabilidad con la que el individuo n elige la alternativa i .

Y_n es el valor de la variable respuesta del individuo n .

C_n es el conjunto de alternativas de n .

V_{ni} es la componente medible de la utilidad de la alternativa i para el individuo n .

V_{nj} es la componente medible de la utilidad de la alternativa j para el individuo n .

En este caso el individuo es cada unidad de carga, mientras que el conjunto de alternativas está formado por dos elementos de transporte: multimodal y carretera.

La unidad de decisión evalúa la función de utilidad de cada alternativa a la que se enfrenta y elige la alternativa con mayor valor de utilidad. Mientras la función de utilidad no se conozca con certeza, ésta es especificada con un término de error aleatorio.

En particular, el MNL para las probabilidades de elección surge si se asume que la utilidad de una alternativa es una función de factores determinantes de la elección de algunos parámetros desconocidos y de la suma de un término de error que sigue una distribución Gumbel. Es esta suposición (que todos los términos de error tengan distribuciones Gumbel y sean estadísticamente independientes) es lo que hace a la formulación del modelo tratable, quedando expresada la utilidad como sigue, tal que la función de utilidad es lineal con respecto al vector de parámetros β^t :

$$U_{nj} = \beta^t X_{nj} + \varepsilon_{nj} \quad j \in C_n,$$

donde:

U_{nj} es la utilidad de la alternativa j para el individuo n .

X_{nj} es el vector de variables explicativas para la alternativa j para el individuo n .

ε_{nj} es el término de error independiente y distribuido Gumbel.

C_n es el conjunto de elecciones para el individuo n .

β^t es el parámetro del modelo.

Además, multiplicando el numerador y el denominador de la ecuación MNL por la cantidad $e^{(-V_{nm})}$ donde V_{nm} es la utilidad medible de una alternativa arbitraria m de un individuo n , se obtiene la siguiente expresión:

$$P_n(i) = \frac{e^{(V_{ni} - V_{nm})}}{1 + \sum_{j \neq m} e^{(V_{nj} - V_{nm})}}, \quad \forall i, j \in C_n.$$

Los MNL se especifican definiendo la utilidad relativa de cada una de las posibles alternativas. Esto implica la definición de las variables explicativas que entran en cada función de utilidad y la relación de los parámetros dentro de la función.

Estas variables pueden ser de varios tipos: características del elemento decisor, que en este caso es el viaje de carga (por ejemplo, el tiempo o el coste de cada viaje) o atributos específicos de cada alternativa (por ejemplo, el tamaño de los vehículos usados en un modo).

Otra distinción de estas variables a considerar se da entre genéricas y específicas. Las genéricas son aquellas que tiene el mismo efecto, es decir el mismo valor del parámetro, mientras que las específicas tienen diferentes efectos en cada alternativa.

En el caso que nos ocupa del presente estudio, los modelos MNL se han aplicado para estimar la probabilidad de que un cargador opte por un medio de transporte u otro en función de la naturaleza de la carga, de las alternativas existentes y de los costes y tiempos asociados a cada alternativa. Por ello, las variables genéricas del modelo son Tiempo y Coste, debido a que generan los mismos efectos mientras que la variable Constante que aparece en la función de utilidad de carretera es específica de este medio de transporte.

En este tipo de métodos las probabilidades de elección varían entre 0 y 1, y la suma de probabilidades de diferentes medios para una unidad de decisión debe ser igual a 1.

3.4.2. Redes Neuronales (RN)

Otra aproximación considerada fue el uso de RN para la elección modal, que es una técnica prometedora para los modelos de elección modal (Hensher and Ton, (2000); Karlaftis and Vlahogianni, (2011); Nijkamp et al., (2004)). Las redes neuronales pueden ajustarse a datos no lineales complejos (Karlaftis y Vlahogianni, (2011), pese a que en general requieren grandes conjuntos de datos, pueden estar sujetas a sobreajustes y los parámetros del modelo no tienen una interpretación directa. Se descartaron en esta memoria porque el tamaño de los conjuntos de datos disponibles era limitado y se deseaba proporcionar una interpretación de los parámetros del modelo.

Las redes neuronales (RN) son estructuras matemáticas, igual que las RB, que permiten representar funciones con varias variables de entrada y de salida inspiradas en el comportamiento de las neuronas, proporcionando un modelo flexible para representar múltiples funciones de elección modal. Pueden ser de una capa o multicapa.

En las RN a las neuronas artificiales se les aplica un conjunto de entradas, multiplicadas por una ponderación equivalente al grado de conexión de la sinapsis. De cada una resulta una salida de otra neurona. Posteriormente, las entradas (ponderadas) se suman y se cuantifica el grado de activación de la neurona. La salida neuronal Y está dada por:

$$Y = f\left(\sum_{i=1}^n (w_i x_i)\right),$$

tal que:

1. la salida neuronal es Y
2. el conjunto de entradas viene dado por x_1, \dots, x_n
3. las ponderaciones de cada entrada son w_1, \dots, w_n .
4. La función de agregación Σ
5. La función de activación f , pudiendo ser ésta una función lineal o no lineal, umbral, que sirve para replicar las transferencias no lineales de las neuronas.

Teorema de Convergencia (Lippman, 1987). Sean $X(n) = \{x_{1(n)}, x_{2(n)}, \dots\}$ las entradas de datos que representan muestras de dos clases linealmente separables, C_1 y C_2 .

Sea w un vector tal que

$$w^T x > 0 \quad \text{si } x \in C_1$$

$$w^T x \leq 0 \quad \text{si } x \in C_2$$

Sea H_1 el subconjunto de entrenamiento que pertenecen a la clase C_1 .

Sea H_2 el subconjunto de entrenamiento que pertenecen a la clase C_2 .

Entonces, si H es un conjunto de entrenamiento linealmente separable, entonces para η positivo, el algoritmo termina (converge).

Es decir, el Perceptrón divide al hiperplano en dos clases siempre y cuando estas sean linealmente separables, tal y como se describe a continuación.

Consideremos el modelo de perceptrón de una capa descrito en la Figura 3.7.

Sea $\theta(n)$ el umbral considerado, equivalente al peso sináptico conectado a una entrada fija e igual a -1. Definimos el vector de entrada como:

$$x(n) = [-1, x_1(n), x_2(n), \dots, x_p(n)]^T,$$

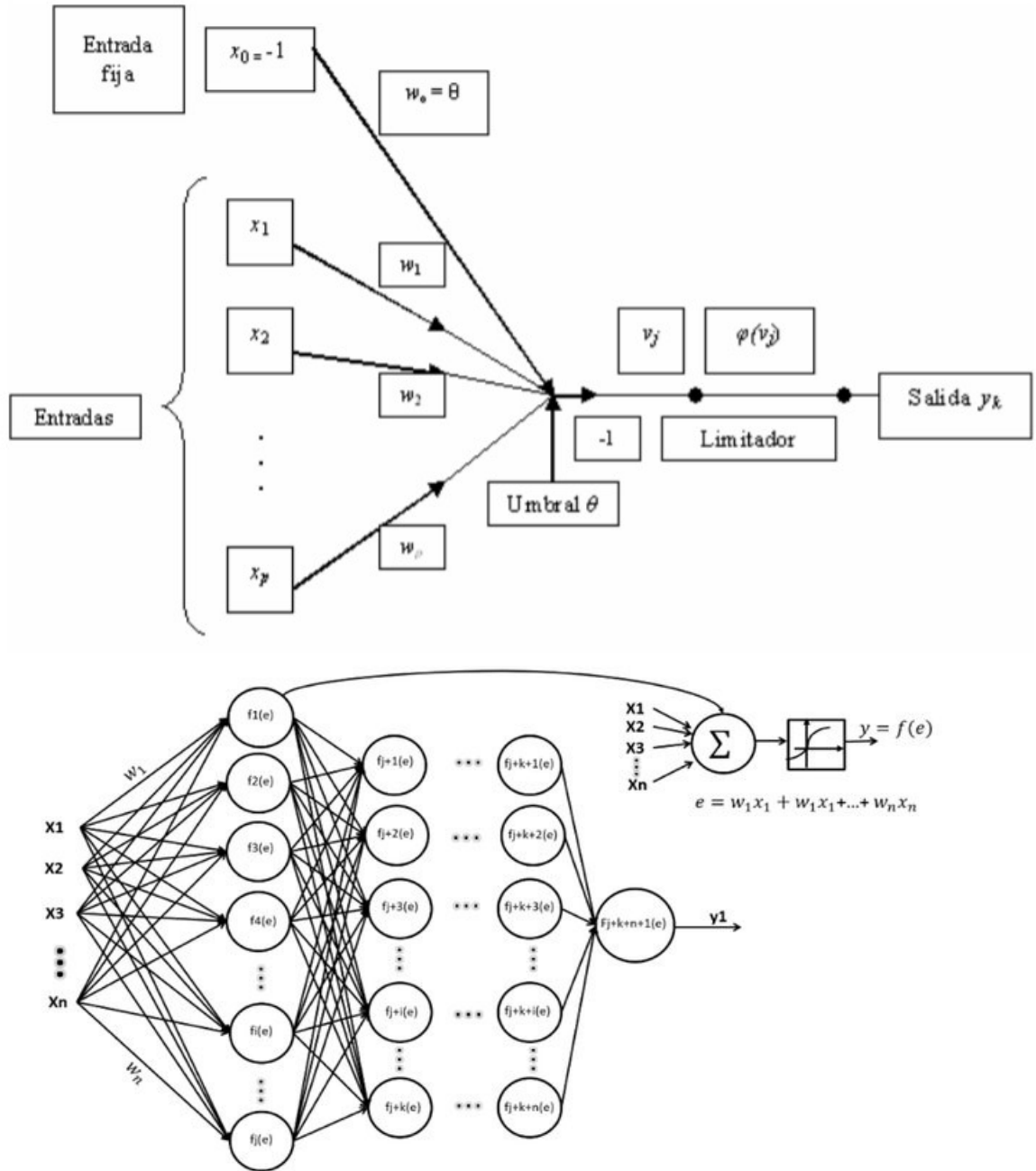


Figura 3.7: Modelo del perceptrón simple con una entrada fija (arriba). Red neuronal multicapa. Fuente: Ramírez, J.A. (abajo)

y el vector de pesos como:

$$w(n) = [\theta(n), w_1(n), w_2(n), \dots, w_P(n)]^T.$$

Entonces, la salida de la combinación lineal es:

$$v(n) = w^T(n) x(n)$$

Para un n fijo, la ecuación de un hiperplano como superficie de decisión de dos clases es $w^T x = 0$, representada en un espacio p -dimensional con coordenadas x_1, x_2, \dots, x_p .

Suponemos que las variables de entrada generan dos clases linealmente separables por un hiperplano, tal que X_1 es el subconjunto de vectores de entrenamiento $x_1(1), x_1(2), \dots$ que pertenecen a la clase C_1 , y X_2 es el subconjunto de vectores de entrenamiento $x_2(1), x_2(2), \dots$ que pertenecen a la clase C_2 .

Dados X_1 y X_2 para entrenar al clasificador, se ajusta el vector de pesos w de forma que las dos clases C_1 y C_2 sean separables. Si existe un w factible entonces las dos clases son linealmente separables. E inversamente, si se sabe que las dos clases son linealmente separables, entonces existe un vector de pesos w tal que cumple con las siguientes expresiones:

$$w^T(n) x(n) > 0 \text{ para cada vector de entrada } x \text{ que pertenezca a la clase } C_1$$

$$w^T(n) x(n) \leq 0 \text{ para cada vector de entrada } x \text{ que pertenezca a la clase } C_2$$

Teorema de Convergencia de incremento finito. *Sean los subconjuntos de entrenamiento X_1 y X_2 linealmente separables y las entradas presentadas al perceptrón las originadas por estos subconjuntos. El perceptrón converge tras n_0 iteraciones, tal que:*

$$w(n_0) = w(n_0 + 1) = w(n_0 + 2) = \dots$$

es un vector solución para $n_0 \leq n_{max}$.

Además, se sabe por el **Teorema de Novikoff (1962)** que no importa lo pequeño que la distancia de separación de la muestral de datos, si ésta es linealmente

separable, entonces el perceptrón encontrará una solución que separa a las dos clases en un número finito de pasos.

El número de iteraciones (tiempos de replicación) dependerá de la distancia de separación y es inversamente proporcional al cuadrado de la distancia.

En el caso de las Redes Neuronales Multicapa, tienen una capa de entrada a la red con su conjunto de neuronas, un número finito de capas ocultas y el conjunto de neuronas de la capa de salida de la red.

Su uso en este caso consiste en estimar las probabilidades de reparto entre alternativas en función de un conjunto de factores de coste y tiempo.

Representan una aproximación de “caja negra” en la que el analista no necesita preocuparse de los detalles internos de cómo funciona la red, aunque debido a esto se pierde capacidad para:

1. Interpretar los resultados del ajuste.
2. Validar si el modelo se comporta de forma razonable.

Una interesante comparación entre RN y modelos Logit es dada por Karlaftis y Vlahogianni, (2011). Estos autores comparan los diferentes términos utilizados en cada aproximación, sus analogías y sus diferencias. En general, las RN están más cerca del campo de la Inteligencia Computacional (CI) o Inteligencia Artificial (IA), que del campo de la Estadística. Una diferencia clave entre ambos enfoques es que las RN actúan como una “caja negra” que no se basa en ningún supuesto sobre la distribución estadística de las variables y que los parámetros del modelo no tienen una interpretación directa, como lo hacen en los modelos Logit. Por otro lado, las RN proporcionan una gran flexibilidad y no están restringidas por los supuestos requeridos para los modelos Logit.

Capítulo 4

Caso de estudio: Corredor Ferroviario Bioceánico Central

4.1. Introducción

Si bien los usuales modelos matemáticos tienen una larga trayectoria y una base sólida, como se ha explicado en capítulos anteriores, a la hora de encontrar soluciones al problema de la elección discreta, existen otro tipo de aproximaciones diferentes a las frecuentistas que pueden dar lugar a mejores resultados, y este es el caso de las aproximaciones bayesianas, en concreto, del uso de las RB aplicado al caso de estudio del Corredor Ferroviario Bioceánico Central (CFBC) que une las costas del océano Atlántico y del océano Pacífico en Sudamérica, atravesando los países de Brasil, Bolivia y Perú.

4.2. Objetivos

4.2.1. Objetivos generales

Los objetivos generales que se han perseguido con este trabajo son: a.1) mejorar el proceso de decisión de viajeros y operadores de carga; a.2) optimizar los modelos de transporte utilizados en planificación; a.3) comprobar que la aplicación de las aproximaciones bayesianas en la etapa de reparto modal supone obtener mejores

resultados; a.4) promover las redes bayesianas en entornos complejos que ya son una realidad como son los Sistemas Inteligentes de Transporte (ITS), los Sistemas Cooperativos (V2V, V2I, V2I2V, I2I) de transporte e infraestructuras, y los vehículos autónomos.

4.2.2. Objetivos específicos

Los objetivos específicos, por otro lado, se resumen en: b.1) conseguir mejores resultados con menor cantidad de información; b.2) disminuir los tiempos de procesamiento para el reparto modal; b.3) mejorar el ajuste coste-beneficio derivado de las mejoras aportadas por las RB en los resultados, cara a optimizar los recursos económicos existentes y conseguir financiación para los proyectos de infraestructura y transporte. En este caso hay que mencionar que el detalle de los costes de cada alternativa de transporte presentada está sujeto a confidencialidad. Sin embargo, en el siguiente enlace oficial se muestra la ficha descriptiva del proyecto correspondiente al COSIPLAN (Consejo Suramericano de Infraestructura y Planeamiento) de UNASUR (Unión de Naciones Suramericanas): http://www.iirsa.org/proyectos/detalle_proyecto.aspx?h=1351. Comprobándose que es uno de los proyectos más importantes del IIRSA (Iniciativa para la Integración de la Infraestructura Regional Suramericana) con una inversión total de unos 7.000 millones de dólares; b.4) mejorar los resultados respecto a los algoritmos utilizados habitualmente.

4.3. Marco de actuación

El modelo descrito se enmarca en el proyecto del Gobierno Plurinacional de Bolivia, financiado por el Banco Interamericano de Desarrollo (BID) con número de préstamo BO-L1056-1, dirigido por el Viceministerio de Transporte (VMT) de Bolivia, a través de su Unidad Técnica Ferroviaria (UTF). Dicho proyecto se denomina “Estudio de Prospectiva Comercial, Mercado y Alternativas Logísticas del Corredor Ferroviario Bioceánico Central (CFBC)”.

Como se ha indicado anteriormente, si bien Bolivia es el país responsable de este estudio, el ámbito del mismo incluye a más países del entorno, como son Brasil y Perú, directamente al ser países cuyas costas se pretende unir, además de otros países



Figura 4.1: Desconexión ferroviaria red oriental y red andina



Figura 4.2: Propuesta de conexión con CFBC

indirectamente, cuyas mercancías y pasajeros también podrían usar este corredor ferroviario, como son Paraguay, Argentina o Chile. La longitud total del corredor ferroviario es de 4.000 km en el que se tienen que rehabilitar algunos de los tramos existentes y construir algunos nuevos (ver desconexión ferroviaria entre la red oriental y la red andina en la imagen siguiente).

Se pretende conectar las redes ferroviarias oriental y occidental de Bolivia, lo cual permitiría un continuo en dicha red uniendo océano Atlántico y Pacífico. Ambas redes bolivianas están separadas por la cordillera de los Andes y esta infraestructura busca conectar el altiplano andino occidental, a 4.000m de altura, y la zona oriental de Bolivia a unos 600m de altura, conectada a su vez a Brasil y Argentina. De esta forma se podrá configurar un corredor ferroviario de costa a costa de gran interés

social y productivo.

Lo que busca es fomentar el desarrollo regional y un efecto integrador, además del intercambio comercial, principalmente entre Bolivia, Perú y Brasil. La nueva construcción se da en la parte central de Bolivia, ya que debido a su orografía la red existente está dividida en dos partes no conectadas, lo que hace que se pierda la conectividad entre ambas costas.

Especialmente, es de suma importancia para las regiones más interiores de Sudamérica, ya que sus comunicaciones con los puertos más cercanos son lentas y costosas, y por ello constituyen un impedimento para el desarrollo económico de la región.

El modelo se empleó para la evaluación de esta infraestructura, es decir, para establecer bajo qué condiciones este corredor va a ser capaz de captar un tráfico de mercancías y pasajeros tal, que haga que sea viable y rentable. Además, permitió el dimensionamiento del mismo, capacidades, vehículos, etc., para alcanzar los objetivos de viabilidad y rentabilidad.

El modelo desarrollado se empleó para dar forma al sistema existente y futuro y plantear los diferentes escenarios de simulación que proporcionaron los valores necesarios para los posteriores análisis. Así a continuación, a partir del modelo conceptual definido con anterioridad, se explicará cada una de etapas tenidas en cuenta y como se han llevado a cabo.

Para la implementación del modelo se empleó la herramienta TransCAD. Este software integra un SIG junto a capacidades de modelización de transporte. Además, cuenta con un lenguaje propio denominado GISDK en el que se ha desarrollado el modelo para su ejecución iterativa, para diferentes periodos de tiempo y diferentes mercancías.

La flexibilidad del modelo permite el establecimiento de tres niveles geográficos de análisis relacionados entre sí: mundial, nacional y regional.

El trabajo descrito en esta memoria parte del modelo de transporte de cuatro etapas que se desarrolló en dicho proyecto por el equipo de expertos nacionales e internacionales, siendo responsable del mismo, para proporcionar estimaciones de los flujos de mercancía y pasajeros que harán uso de dicho corredor frente a las alternativas existentes. Para mercancía, los principales modos en competencia con el

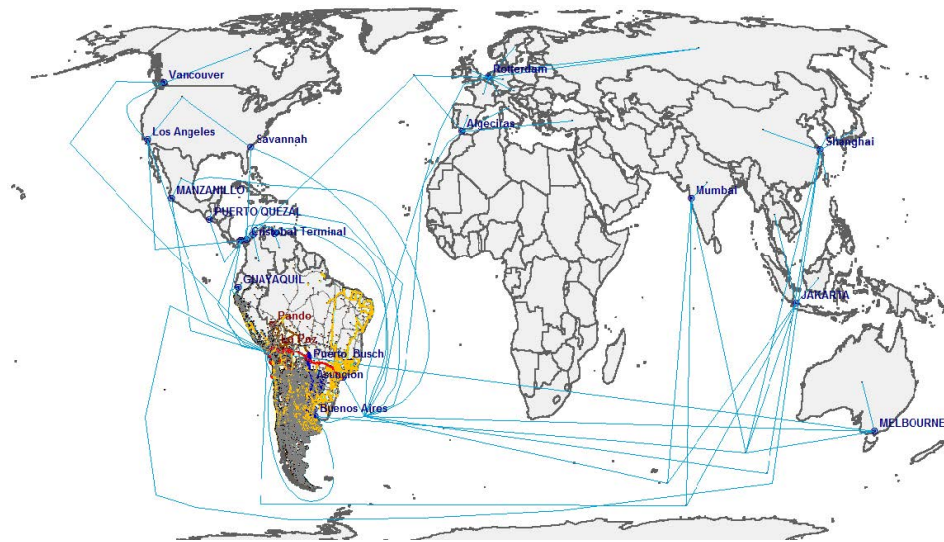


Figura 4.3: Modelo de transportes (mundial)

ferroviario, por agregación en la cadena de transporte, son la carretera y la hidrovía Paraguay-Paraná. Una descripción completa puede leerse en Rios-Prado et al. (2013), si bien esta memoria se centra en la problemática del transporte de mercancías ya que es el de mayor interés para el CFBC.

Tomando el esquema general, explicado en el modelo conceptual, la Figura 3.4 representa los pasos que se llevaron a cabo en la creación del modelo de transportes del CFBC y la siguiente figura muestra el ámbito mundial modelizado a tales efectos.

4.4. Metodología

Recoger todas las citas y documentación consultada es prácticamente imposible, por lo que mencionar a modo de muestra los trabajos recientes de Sun et al. (2006), Correa et al. (2009), Tang et al. (2012), aplicando RB a modelos de transporte de vehículo privado y pasajeros, **Daziano et al. (2013)**, donde están contenidas las RB asociadas a la distribución multinomial, **Yannis Tyrinopoulos (2013)** y **Tai Yu Ma (2015)** donde pueden verse distribuciones más generales aplicadas a este tipo de redes.

Sin embargo, todos ellos se focalizan en ámbitos de estudio urbanos o metro-

politanos y de vehículo privado, bien para analizar temas como; la seguridad vial, considerando los beneficios derivados del uso de modelos Markov Chain Monte Carlo (MCMC) y de los modelos bayesianos Empirical Bayes (EB) y Full Hierarchical Bayes (FHB); la movilidad en situaciones de congestión vehicular, con modelos Probit; o los patrones de movilidad urbana de los trabajadores (commuters y similares) de zonas frontera metropolitanas, analizando cuáles son las principales variables que influyen en un modelo de reparto modal orientado a la promoción del transporte público en detrimento del vehículo privado. Por tanto, bajo esta consideración, si bien son similares en cuanto a la orientación de la elección modal, difieren en cuanto a su uso hacia un modelo de transporte de mercancías como el indicado en el caso de estudio que se presenta, el cual además es un caso real.

La metodología utilizada responde a las siguientes preguntas clave del desarrollo del trabajo.

¿Cuáles son las contribuciones originales? A la vista de la base teórica explicada en los capítulos anteriores sobre Redes Bayesianas y Modelos de Transporte, la principal novedad reside en el uso de las redes bayesianas en la etapa de reparto modal de un modelo de transportes (de cuatro etapas) de mercancías, multimodal, a nivel mundial, confirmando que los resultados obtenidos son mejores que los obtenidos con otras técnicas más habituales, como los modelos multinomial logit (MNL) y las redes neuronales (RN).

¿Cuál es su naturaleza? La definición de la temática y el planteamiento del problema de investigación centran el estudio en la planificación de las infraestructuras a futuro. En este caso concreto, de una infraestructura multimodal con relaciones a todos los niveles posibles: distritales, regionales, estatales y continentales. El desarrollo de la perspectiva teórica se basa en las Redes Bayesianas, en la definición, tipología y aplicación de las mismas, considerando sus principales características y propiedades. La recolección, análisis e interpretación de los datos toma de referencia un caso práctico de estudio, un corredor ferroviario que una el océano Atlántico con el océano Pacífico, prioritariamente para el transporte de mercancías. Siendo las variables principales en la elección de las alternativas de transporte multimodal las de tiempo y coste, la naturaleza de la investigación es meramente cuantitativa. Si bien, la posibilidad de incluir otro tipo de variables cualitativas hace más extensa

esta naturaleza inicial. La estrategia metodológica se basa en la comparativa de los resultados obtenidos para las distintas alternativas de transporte, en cuanto a la elección del modo de transporte (reparto modal), en el modelo de transportes construido a tal efecto. Las metas del estudio, finalmente, han sido las de comprobar que los resultados obtenidos con las redes bayesianas son mejores que las obtenidas con otros métodos habituales, los multinomial logit (MNL) y las redes neuronales (RN).

¿Cuál es su alcance? El alcance del uso de las RB es: a) macroscópico, como es el caso de estudio presentado, pudiéndose aplicar de forma añadida tanto a la etapa (1) de generación-atracción, como a la etapa (2) de distribución de viajes (transformada de la matriz de generación-atracción a la matriz de origen-destino mediante una función de impedancia). El siguiente paso, lógico, de la presente investigación precisamente va en línea con la implementación de este tipo de redes bayesianas en tres de las cuatro etapas de un modelo de mercancías; b) microscópico, en los modelos de tráfico de ámbito más urbano o interurbano, pudiéndose aplicar a los modelos de elección de ruta o seguimiento vehicular, inicialmente.

¿Cuáles son sus limitaciones? Hasta el momento: a) la inexistencia de proyectos similares de transporte de mercancías en el que se hayan podido replicar el uso de redes bayesianas al reparto modal, que verifiquen que los resultados obtenidos en este caso; b) la difusión de las redes bayesianas como herramientas de aplicación en los proyectos de planificación de infraestructura todavía tiene un margen de penetración en las entidades públicas y empresas de ingeniería, por lo que a día de hoy su uso no está muy extendido.

Inicialmente, se opta por el empleo de un modelo de elección modal (MEM) o de reparto modal, que estima las proporciones de uso de las alternativas correspondientes a combinaciones de modo de transporte y puerto, empleando como factores los tiempos y costes de cada alternativa, el puerto marítimo en cuestión, y ajusta un modelo por cada tipo de mercancía.

Como alternativa añadida se presenta una aproximación bayesiana al problema de la elección modal, que no ha sido considerada por autores previos en este tipo de problemas de modelización de transporte. La red bayesiana permite estimar las proporciones de uso de las alternativas por sus tiempos y costes y añadiendo también la variable precio FOB (free On Board) del producto. En el modelo MEM no se usa

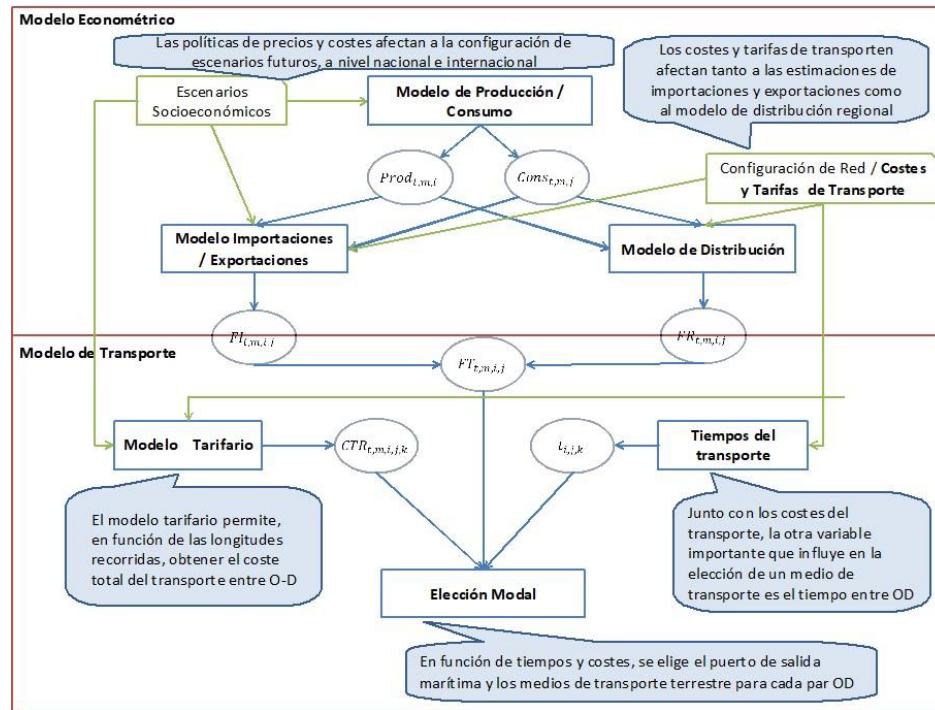


Figura 4.4: Esquema relacional del modelo de transportes

el precio FOB del producto, sino que se ajusta un modelo independiente para cada producto.

El precio FOB es el valor de la mercancía a bordo de un transporte marítimo, el cual incluye: costo de la mercancía en el país de origen, transporte de los bienes y derechos de exportación. Este valor está relacionado con el uso del Incoterm FOB.

El Incoterm Free On Board (FOB), que se traduce como ‘franco a bordo’ se utiliza exclusivamente para transporte marítimo o fluvial, tal que a efectos comerciales internacionales:

1. El vendedor entrega la mercancía en el puerto de embarque y asume los costos de trámites aduaneros de exportación y licencias de exportación.
2. El comprador realiza los trámites de importación, consigue el transporte desde el puerto de embarque y asume los costos durante la entrega de la mercancía (descarga, flete, despacho, etc.)

En la Figura 4.4 se muestra el esquema relacional del modelo de transportes

construido sobre el que se aplicará la aproximación bayesiana al modelo de reparto o elección modal.

4.5. Base de datos

Para realizar el ajuste de los modelos de elección, se partió de la mayor muestra disponible que permite estimar la distribución geográfica de los movimientos de carga en Bolivia, que es la base de datos del Instituto Nacional de Estadística (INE) boliviano para Comercio Exterior.

Dicha base de datos contiene registros de exportaciones e importaciones identificando las siguientes variables:

1. Productos agrupados por diversas categorías (se han empleado las categorías de sección NANDICA de cuatro dígitos y la clasificación de principales productos).
2. Departamento de origen o destino.
3. País de origen o destino de importación o exportación.
4. Vía de salida, que permite identificar a través de qué puerto se envía la carga en aquellos envíos a destinos internacionales de larga distancia.
5. Modo de salida, que permite identificar si se está empleando carretera, ferrocarril o hidrovía.

Esta información en conjunto con el modelo de costes y tiempos, obtenidos de la información oficial disponible en el Viceministerio de Transporte de operadores de transporte y carga bolivianos, implementado con el software de modelización macroscópica TransCAD (Caliper Corp., www.caliper.com), permite analizar los modos de salida adoptados en relación con los costes y tiempos que implica cada opción.

Para ello se combinaron tablas de importaciones y exportaciones descargadas del INE boliviano con las matrices OD de costes y tiempos obtenidas con el modelo de transporte. La tabla de datos de exportaciones empleada para el ajuste de los modelos de elección contiene un total de 466 registros y la de importaciones 4.333.

Los principales datos de entrada al modelo se resumen en:

1. Cartografía oficial del VMT de la red de transporte considerando los modos de transporte carretera, ferrocarril, mar e hidrovía Paraguay-Paraná. Esto permite construir el grafo de arcos y nodos de la red.
2. Bases de datos del Instituto Nacional de Estadística de Bolivia (INE).
3. Producciones agrícolas, la superficie cultivada y la productividad del terreno desde 1990 hasta 2011.
4. Producciones mineras en el mismo periodo.
5. Exportaciones e importaciones, para el periodo 2003-2012.
6. Índices de volumen físico de producción y cantidad de transporte.
7. Producciones y consumos de los productos generales. Basado en informes sectoriales, de productores y variables socioeconómicas.
8. Estimaciones de capacidades futuras, reservas minerales existentes y extensiones agrícolas.
9. Informes sectoriales y de encuestas.
10. Proyecciones económicas de la OCDE y organismos internacionales.

4.6. Alternativas de elección

El problema de elección modal en este caso de estudio considera como alternativas de transporte las combinaciones de puertos de importación/exportación, junto con los modos de transporte hasta el puerto. No todas las combinaciones entre opciones son viables en la práctica, por lo que un primer paso consistió en identificar las combinaciones de alternativas factibles.

Las condiciones de partida respecto de las alternativas de transporte disponible son: las opciones relativas al uso de la hidrovía se integran con el modo marítimo, de esta manera no se hace distinción entre puertos fluviales y marítimos; no todos los tipos de mercancía hacen uso de todos los puertos. Como alternativas de transporte

Cuadro 4.1: Alternativas disponibles en el modelo (C=carretera, MM=multimodal carretera/tren). * Esta alternativa no está disponible en la actualidad, sino que será fruto de la construcción del CFBC.

Alternativas			
Puerto	Modo	Graneles	Contenedores
Arica	C	Sí	Sí
Arica	MM	No	No
Ilo	C	Sí	Sí
Ilo	MM*	Sí	Sí
Iquique	C	Sí	Sí
Antofagasta	C	Sí	No
Antofagasta	MM	Sí	No
Pto. Busch	C	Sí	Sí
Pto. Busch	MM	Sí	Sí
Pto. Suarez	C	Sí	Sí
Pto. Suarez	MM	Sí	Sí
Buenos Aires	C	Sí	Sí
Santos	C	Sí	Sí
Santos	MM	Sí	Sí

para el modelo de elección se consideran los pares de puerto/modo siendo los modos sólo carretera y multimodal carretera / ferrocarril.

La Tabla 4.1 muestra las alternativas disponibles, combinación de puerto y modo, para las mercancías de tipo graneles y para la mercancía contenerizada.

4.7. Ficheros principales de datos y estructura

La información de partida se almacena, depura, valida y normaliza para su posterior tratamiento en cada uno de los modelos considerados: Multi Nomial Logit, Red Neuronal y Red Bayesiana.

elecmod.txt: Bloc de notas

DIR	FREIGHT ID1	ID2	FOBPrice	MTPA	T1	T2	T3	T4	T5
1	EXP	Contenedorizado	17736	17570	-0.767612711	-0.9976888083	-0.603929412	-0.54	-0.54
2	EXP	Contenedorizado	17736	17571	-0.777452525	-0.975557663	-0.603929412	-0.54	-0.54
3	EXP	Contenedorizado	17736	17572	-0.780937486	-0.992979908	-0.603929412	-0.54	-0.54
4	EXP	Contenedorizado	17736	17574	-0.780938928	-0.999962171	-0.603929412	-0.54	-0.54
5	EXP	Contenedorizado	17736	17575	-0.78246522	-0.926703606	-0.603929412	-0.54	-0.54
6	EXP	Contenedorizado	17736	17576	-0.764813398	-0.998700375	-0.478023529	-0.41	-0.41
7	EXP	Contenedorizado	17736	17577	-0.780072907	-0.99896473	-0.603929412	-0.54	-0.54
8	EXP	Contenedorizado	17736	17579	-0.77806839	-0.999621576	-0.603929412	-0.54	-0.54
9	EXP	Contenedorizado	17736	17580	-0.782711951	-0.983363093	-0.603929412	-0.54	-0.54
10	EXP	Contenedorizado	17736	17581	-0.781622354	-0.913080335	-0.603929412	-0.54	-0.54
11	EXP	Contenedorizado	17736	17582	-0.783661824	-0.976201214	-0.603929412	-0.54	-0.54
12	EXP	Contenedorizado	17736	17583	-0.782117871	-0.999860495	-0.279011765	-0.21	-0.21
13	EXP	Contenedorizado	17736	17584	-0.783780877	-0.997510578	-0.279011765	-0.21	-0.21
14	EXP	Contenedorizado	17736	17585	-0.781640966	-0.999835259	-0.137247059	-0.07	-0.07

Figura 4.8: Fichero de texto elecmod

Los ficheros principales de datos que se utilizan como entrada se denominan:

1. Dataset Elección Modal (libro Excel)
2. RDataset (libro Excel)
3. Elecmod (archivo txt)

4.7.1. Archivo Excel Dataset Eleccion Modal

Contiene los datos iniciales junto las transformaciones realizadas para la RB y la RN.

1. Hoja DATASET CFBC original: datos de partida de comercio exterior (INE).
2. Hoja DATASET Eleccion Modal: datos transformados para usar en R (inputs).
3. Consulta Análisis: consulta dinámica al *dataset* inicial para agrupar los registros por OD & tipos de mercancía, cuyas columnas corresponden a los distintos modos de transporte.

4.7.2. Hoja Excel DATASET Elección Modal

A continuación se describen las tablas del DATASET Elección modal resumido en la Figura 4.5. De arriba a abajo la forma de interpretarlo es la siguiente:

Las primeras columnas muestran el sentido de Import/Export, el tipo de mercancía, los ID de origen y destino, el precio FOB de la mercancía (\$/ton), y el flujo de ton/año.

Las siguientes columnas contienen, para cada alternativa de transporte los flujos de carga en toneladas/año (ton/año).

A partir de las ton/año de cada alternativa se determinan los porcentajes de reparto observados reales entre alternativas. Este valor es fundamental ya que determina las probabilidades de uso reales por sentido import/export, tipo de mercancía y par OD, para cada una de las combinaciones ID1 e ID2. El resultado es el porcentaje obtenido de la división de toneladas/año para cada combinación de las anteriormente

indicadas, entre la suma total de carga por OD. En base a estos porcentajes, posteriormente se determinan los errores relativos y absolutos del Multi Nomial Logit, Red Neuronal y Red Bayesiana, lo que mostrará la bondad en el ajuste de cada una de las técnicas utilizadas.

Posteriormente, se muestran los tiempos de viaje (horas) de cada registro por alternativa de transporte. Las celdas con títulos en azul son las alternativas de carretera y, en naranja alternativas de ferrocarril.

De la misma forma, después, se tienen los costes en \$/ton.

En las siguientes dos columnas se determinan el mínimo tiempo y coste de entre todas las alternativas.

Finalmente, a partir de las diferencias con los tiempos y costes mínimos de cada registro, por OD, en la última tabla se calculan los Incrementos de Tiempo y Coste respecto a la mejor alternativa.

Por ejemplo, en la tabla siguiente la columna etiquetada como *T_ARICA_Ferrocarril* indica que para ese registro el menor tiempo de viaje correspondería a la opción de acceder al puerto de Arica por Ferrocarril (igual a 0), ya que el valor de acceder a Arica por Carretera es mayor (igual a 118). Todos los tiempos de ese registro corresponden a la diferencia en tiempo con respecto al mejor caso.

Con respecto a los costes la interpretación es similar.

4.7.3. Archivo Excel RDataset

Este archivo contiene las siguientes hojas de cálculo:

1. **DATA1**: datos procedentes del fichero Excel “Dataset Eleccion Modal”, explicado anteriormente, normalizados al rango (-50, 800) en el caso de los tiempos, y al rango (-50, 500) en el caso de los costes. Este fichero contiene los siguientes grupos de columnas en orden de izquierda a derecha:
 - a) Encabezados gris: tipos de mercancía / ID por OD, precio FOB, TMPA.
 - b) Encabezados azul: incremento en tiempo respecto a la mejor alternativa.
 - c) Encabezados naranja: Incremento en coste respecto a la mejor alternativa.

- d) Encabezados azul: incremento en tiempo respecto a la mejor alternativa, normalizado al rango (-50, 800).
 - e) Encabezados naranja. incremento en coste respecto a la mejor alternativa normalizado al rango -50, 500.
 - f) Encabezado azul oscuro. Porcentajes de reparto modal (valor fundamental para realizar la comparativa entre los distintos métodos MNL, RN y RB).
2. **LOGIT**: incluye los datos anteriores y los modelos Logit ajustados del proyecto CFBC, de tal forma que muestra los modelos propios del mismo para estimar el reparto de mercancías dado por los Logit. Estos datos sirven para luego compararlos con los resultados de las RN y RB.
 3. **DATANORM**: conjunto de datos normalizados para la exportación a R, así como una serie de columnas para construir la cadena de caracteres del modelo usado en R.
 4. **Resultados RB y Resultados RRNN**: contienen los resultados obtenidos de la programación en R.
 5. **RESUMEN**: contiene la comparación entre modelos.

4.7.4. Fichero de texto elecmod

Este fichero contiene los datos del Excel exportados en formato compatible con R. Además, dicho fichero puede ser importado en RStudio con la opción **Import Dataset**.

4.8. Propósitos del modelo de elección modal

Distintos cargadores toman decisiones sobre la elección entre las alternativas de transporte disponibles, que dependen de muchos factores particulares los cuales no se pueden recoger en el modelo general por la tipología de información.

Por tanto, los propósitos principales del modelo de elección modal son:

1. Reflejar cómo, dados unos factores sobre los que interesa estudiar variaciones, se reparten los flujos de carga entre las opciones disponibles.
2. Estimar las elasticidades: cuánto varían los porcentajes de reparto al cambiar los factores de decisión.
3. Estudiar los dos principales factores de la elección modal:
4. **Tiempo.** Para saber cómo afecta la reducción de tiempos de viaje que permitirá lograr el CFBC, una vez en operación.
5. **Coste.** Para saber cómo los precios del CFBC afectarán a las captaciones de mercancía.

4.9. Calibración del MNL

Se llevaron a cabo con el software estadístico R por ser la herramienta estadística con mayor variedad de librerías de programación y técnicas de este tipo desarrolladas por investigadores, teniendo en cuenta además las librerías disponibles específicas para RB mencionadas en el capítulo correspondiente.

Las funciones de utilidad de las mercancías (exportadas / importadas), se distinguieron en:

1. Flujos con salida al mar:

$$\begin{aligned}
 U_{i,j,m,t,carretera,p} = & \\
 & \beta_0 + \\
 & \beta_1(\tau i/p, p/j, m, t, carretera + \tau p/i, j/p, m, t, mar) + \\
 & \beta_2(ci/p, p/j, m, t, carretera + cp/i, j/p, m, t, mar)
 \end{aligned}$$

$$\begin{aligned}
 U_{i,j,m,t,multi,p} = & \\
 & \beta_0 + \\
 & \beta_1(\tau i/p, p/j, m, t, multi + \tau p/i, j/p, m, t, mar) + \\
 & \beta_2(ci/p, p/j, m, t, multi + cp/i, j/p, m, t, mar)
 \end{aligned}$$

2. Flujos internos sin salida al mar:

$$\begin{aligned}
 U_{i,j,m,t,carretera,p} = & \\
 & \beta_0 + \\
 & \beta_1(\tau_{i,j,m,t,carretera}) + \\
 & \beta_2(c_{i,j,m,t,carretera})
 \end{aligned}$$

$$\begin{aligned}
 U_{i,j,m,t,multi,p} = & \\
 & \beta_0 + \\
 & \beta_1(\tau_{i,j,m,t,multi}) + \\
 & \beta_2(c_{i,j,m,t,multi})
 \end{aligned}$$

siendo

$\tau_{i/p,p/j,m,t}$: tiempo entre nodo de salida (exportación) / llegada (importación) y puerto de la ruta. Carretera / Multimodal carretera y ferrocarril. Por OD, periodo temporal y tipo de carga.

$\tau_{p/i,j/p,m,t,mar}$: tiempo entre nodo de salida / llegada y puerto de la ruta, por mar. Por OD, periodo temporal y tipo de carga.

$c_{i/p,p/j,m,t}$: coste entre nodo de salida / llegada y puerto de la ruta. Carretera / Multimodal carretera y ferrocarril. Por OD, periodo temporal y tipo de carga.

$c_{p/i,j/p,m,t,mar}$: Coste entre nodo de salida / llegada y puerto de la ruta, por mar. Por OD, periodo temporal y tipo de carga.

$\tau_{i,j,m,t}$: Tiempo de viaje entre par OD. Carretera / Multimodal.

$c_{i,j,m,t}$: coste de viaje entre par OD interior. Carretera / Multimodal.

$\tau_{i,j,m,t}$: tiempo de viaje entre par OD. Carretera / Multimodal.

i : ID de zona de transporte de origen.

j : ID de zona de transporte de destino.

p : ID de puerto.

m : ID de mercancía.

t : D de periodo temporal.

multi: ruta multimodal.

Las consideraciones iniciales del modelo fueron: ajuste por cada tipo de mercancía (contenedores, graneles limpios, sucios y líquidos); parámetros asociados a valores de tiempos y costes de cada alternativa; parámetros para recoger la preferencia por modos o puertos como en el caso de los contenedores (se había observado una preferencia por el puerto de Ilo a igualdad de condiciones de tiempo y coste debido a sus instalaciones portuarias); se incluye un parámetro dependiente alternativo para el puerto de Arica que incrementaba su atractividad como entrada/salida al mar, debido a las instalaciones de dicho puerto para el transporte de contenedores; el puerto de Santos, es menos atractivo para la mercancía boliviana debido a la distancia y a la congestión del mismo.

De tal forma que así se consideraron tres tipos de parámetros en el modelo:

1. Un parámetro de efectos fijos asociado al puerto.
2. Un parámetro de efecto del coste por cada tipo de mercancía.
3. Un parámetro de efecto del tiempo por cada tipo de mercancía.

Los valores de los parámetros ajustados se omiten por estar sujetos a cláusulas de confidencialidad del proyecto del Viceministerio de Transporte boliviano, pero en la tabla siguiente puede comprobarse la significación de los parámetros de los modelos, en este caso para el Puerto de Arica, que es una de las alternativas portuarias más importantes en cuanto al transporte de mercancías, del total de alternativas existente para la exportación e importación.

Los parámetros de la función de utilidad que no mostraron ser estadísticamente significativos en el test de significación fueron eliminados del modelo hasta seleccionar los modelos indicados en la Tabla 4.2.

Cuadro 4.2: Parámetros función de utilidad / p-valor del test de significación

Tipo de Carga	Parámetro	p-valor	Significación
Contenedores	Preferencia Puerto Arica	0.0000	SI
Contenedores	Coste	0.0000	SI
Contenedores	Tiempo	0.0000	SI
Graneles Sólidos Sucios	Coste	0.0000	SI
Graneles Sólidos Sucios	Tiempo	0.0000	SI
Graneles Sólidos Limpios	Coste	0.5331	NO
Graneles Sólidos Limpios	Tiempo	0.0001	SI
Graneles Líquidos Sucios (combustibles)	Coste	0.0000	SI

La Tabla 4.2 muestra el ajuste resultante con la información utilizada en situación base, en el que se verifica que las variables coste y tiempo son determinantes en la elección modal, considerando las funciones de utilidad por tipo de mercancía, a excepción del coste en los graneles sólidos limpios.

Así también, la variable auxiliar correspondiente al efecto de la preferencia por el puerto de Arica en las alternativas de transporte analizadas resultó ser estadísticamente significativa, lo que equivale a interpretar que los cargadores, **a igualdad de condiciones**, prefieren usar este puerto debido a las buenas condiciones que tiene para las operaciones portuarias.

Esto se traduce, como veremos en las siguientes imágenes, en que este puerto se sitúa como el óptimo de las alternativas posibles de conexión por carretera para la exportación de entre las opciones de conexión en situación base. Ese hecho encaja con la realidad al ofrecer instalaciones especialmente aptas para el tráfico de contenedores que no se explican sólo en término de tiempos y costes de acarreo.

En el caso de los graneles sólidos limpios, como se ha adelantado antes, sólo la variable tiempo resultó ser significativa, lógico al ser principalmente de productos agrícolas para los que los tiempos de viaje son determinantes ante posibles deterioros o perecimientos. Para los graneles líquidos (principalmente combustible) el principal factor de decisión se observó que era el coste, ya que el tiempo no influye en su deterioro.

Así, los resultados del MNL se utilizaron para estimar la división de los flujos para las alternativas terrestres, por tipo de carga y a nivel global. De esta forma se obtienen las matrices OD de flujos por alternativas α , $F\alpha, t, m, i, j$, tal que es factible conocer las captaciones de carga por modo, en base a las utilidades (probabilidades de uso de las alternativas de transporte) y las asignaciones a red relativas al reparto modal estimado.

En la Tabla 4.3 se muestra un ejemplo asignación a red, todo o nada, con el flujo (toneladas de carga transportada). Los modos de transporte se representan así: gris=carretera, rojo=ferrocarril, azul claro=marítimo y azul oscuro=fluvial (hidrovía Paraguay-Paraná). El grosor de cada modo representa las toneladas transportadas en cada tramo (total flow). La imagen superior muestra el flujo de carga en la situación base y en la imagen inferior en la situación futura, con el corredor ferroviario funcionando al 100%. Por motivos de confidencialidad no se muestran las cantidades transportadas, tipo de mercancía y año de análisis.

Las conclusiones derivadas de los resultados de los errores de estimación de las probabilidades de uso respecto a los datos observados reales son:

1. Principalmente, el valor del error medio cuadrado y el error máximo absoluto para la conexión con el puerto de Arica por carretera son muy elevados. Esto es peligroso ya que se ha demostrado que este puerto es una de las salidas más importantes de las exportaciones hacia el océano Pacífico. Por tanto, desviaciones así repercuten en un mal dimensionamiento de los recursos e infraestructuras, y supone una mala planificación a futuro y unos condicionantes económicos y financieros de inversión mucho más elevados a largo plazo .
2. En relación a la otra alternativa principal de salida, el puerto de Antofagasta, el error medio cuadrado obtenido está en unos valores razonables, si bien, los errores máximos absolutos son muy elevado, entre un 93 % y un 99,5 %, lo cual es peligros para las estimaciones futuras.
3. Además, se suma el desconocimiento de las relaciones existentes de dependencia entre las variables significativas, que sí existen en las Redes Bayesianas.

Las conclusiones de los anteriores gráficos a partir de las estimaciones del modelo

Cuadro 4.3: Resultados ajuste MNL

	Error Medio Cuadrado (MNLogit)	Error Máximo Absoluto (MNLogit)
ANTOFAGASTA_Road (carretera)	15,08 %	99,48 %
ANTOFAGASTA_Rail (ff.cc.)	16,26 %	93,73 %
ARICA_Road (carretera)	73,71 %	100,00 %
ARICA_Rail (ff.cc.)	19,67 %	41,66 %
ILO_Road (carretera)	14,05 %	100,00 %
IQUIQUE_Road (carretera)	27,10 %	100,00 %
IQUIQUE_Rail (ff.cc.)	5,64 %	10,24 %
NUEVAPALMIRA_Road (carretera)	19,56 %	100,00 %
PTOSUAREZ_Road (carretera)	6,84 %	95,49 %
PTOSUAREZ_Rail (ff.cc.)	5,03 %	90,26 %
SANTOS_Road (carretera)	8,28 %	100,00 %
SANTOS_Rail (ff.cc.)	3,78 %	10,24 %

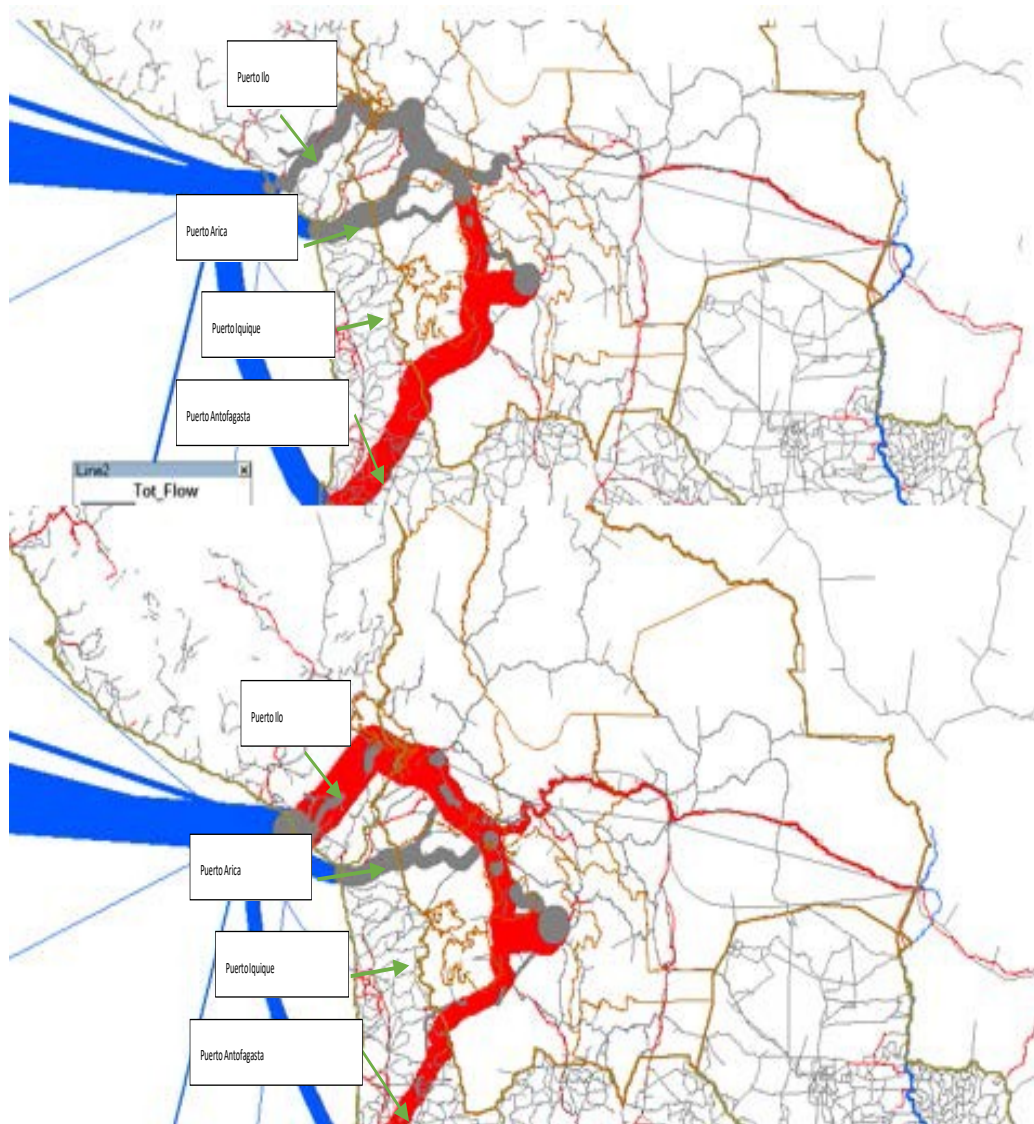


Figura 4.9: Ejemplo de comparativa situación base (arriba) y situación futura (abajo)

obtenidas en la anterior tabla, tomando de ejemplo las exportaciones bolivianas por el océano Pacífico:

1. Situación base: la vía principal de salida de la carga es la conexión hasta el puerto de Antofagasta y desde aquí por vía marítima. Hasta este puerto el tráfico llega en modo ferroviario. También se realizan salidas, pero menores, por Arica e Ilo, conectando por carretera.
2. Situación futura: con el corredor ferroviario funcionando al 100 %, éste se posiciona como un modo competitivo frente a la carretera, en costes y tiempos, para el transporte de mercancías. De esta forma, las mercancías se reparten entre Antofagasta y el puerto de Ilo. En este último caso se aprecia como el flujo de carga es mayor por ferrocarril que por carretera, con lo que el total de carga transportada aumenta respecto la situación base.

Por tanto, las funciones de utilidad resultantes de la aplicación del MNL para calcular la elección modal son consistentes en la situación base, y por tanto, pueden ser usadas para las proyecciones de carga en situación futura. Esto es importante en la etapa de asignación a la hora de calcular las captaciones posibles en el corredor ferroviario.

4.10. Calibración de la RN

En el caso de la Red Neuronal, los datos de partida se muestran en la Tabla 4.4. En cuanto a los rangos de las variables de entrada, son:

1. Precio FOB del producto. Normalizada al rango -1 a +1.
2. Flujo anual en toneladas métricas (MTPA). Normalizada al rango -1 a +1.
3. Diferencia de tiempo de cada alternativa con respecto a la alternativa de tiempo mínimo (T1 a T12). Normalizada al rango -1 a +1.
4. Diferencia de costes de cada alternativa con respecto a la alternativa de coste mínimo (C1 a C12). Normalizada al rango -1 a +1.

Cuadro 4.4: Variables de entrada de la RN

Input	Variable	Description
Time	$T_k = t_k - \min_k t_k$	Time difference to the best alternative
Cost	$C_k = c_k - \min_k c_k$	Cost difference to the best alternative
Freight flow per year	$F_{i,j,m}$	Freight flow from the Origin to the Destination.
FOB Price	$FOB_{i,m}$	Free on Board price of the product at the Origin

	DIR	FREIGHT	ID1	ID2	FOBPrice	MTPA	T_ARICA	CART	IQUIQUE	CT	PTOSUARE	T_ILO	CARRE	T_NUEVAPAL
1	EXP	Contenedoriz	17736	17570	7.66219443	377,908.0	118	145	75	134	335			
2	EXP	Contenedoriz	17736	17571	3.10352881	4,004,938.9	118	145	70	134	330			
3	EXP	Contenedoriz	17736	17572	1.48898919	1,149,547.1	118	145	209	134	468			
4	EXP	Contenedoriz	17736	17574	1.48832083	5,200.0	118	145	249	134	509			

Figura 4.10: Ejemplo de muestra de datos de exportaciones e importaciones

En el caso de las variables de salida, el porcentaje de elección de cada alternativa está Normalizada al rango -0,9 a 0,9.

Es decir, todas las variables de entrada de la red neuronal están normalizadas en el rango de -1 a +1, y las salidas de la red neuronal son las probabilidades de elección modal de cada alternativa de transporte p_k . Por lo tanto, hay una salida para cada alternativa k que se evalúa para cada par origen-destino y tipo de flete resultantes de la elección modal $p_{m,i,j,k}$.

En la Figura 4.10 se muestra una imagen del conjunto de datos utilizados de exportaciones e importaciones.

El tipo de red neuronal elegida para este caso fue un perceptrón multicapa con una capa oculta equipada con el paquete de red neuronal del software R. La función de activación de las neuronas es una función logística. Todos los pesos entre neuronas se ajustan mediante un algoritmo de entrenamiento por retropropagación de errores.

El entrenamiento se realizado con la librería *neuralnet* implementada en R y se utiliza el algoritmo de entrenamiento de retropropagación. Riedmiller et al. (1993). La red ajustada tiene 2 neuronas en la capa oculta, ya que al probar con 4 y 6 neuronas en la capa oculta no se observan mejoras significativas.

Cuadro 4.5: Resultados ajuste RN

	Error Medio Cuadrado (RN)	Error Máximo Absoluto (RN)
ANTOFAGASTA_Road (carretera)	14,389 %	92,214 %
ANTOFAGASTA_Rail (ff.cc.)	16,262 %	98,735 %
ARICA_Road (carretera)	30,690 %	97,492 %
ARICA_Rail (ff.cc.)	0,034 %	0,730 %
ILO_Road (carretera)	13,723 %	96,980 %
IQUIQUE_Road (carretera)	24,377 %	91,559 %
IQUIQUE_Rail (ff.cc.)	0,004 %	0,091 %
NUEVAPALMIRA_Road (carretera)	14,497 %	100,000 %
PTOSUAREZ_Road (carretera)	5,254 %	99,959 %
PTOSUAREZ_Rail (ff.cc.)	4,201 %	88,618 %
SANTOS_Road (carretera)	8,092 %	99,786 %
SANTOS_Rail (ff.cc.)	0,010 %	0,178 %

En la Figura 4.11 se muestra el gráfico de la red neuronal entrenada.

La red neuronal tiene un total de 30 entradas correspondientes al precio FOB, el flujo de carga por año en toneladas y la diferencia de tiempo y costo para cada una de las 14 alternativas de transporte. La red tiene solo dos neuronas en la capa oculta. Se entrena con el 80 % de los registros del conjunto de datos, mientras que el 20 % restante se utiliza para la validación cruzada.

En la Tabla 4.5 se muestran los errores obtenidos de la estimación con la RN, para cada una de las alternativas analizadas.

Las principales conclusiones con este ajuste son, a efectos de error respecto a la elección modal de los datos observados reales, considerando que no se han utilizado estos datos para realizar ninguna asignación a red ya que no es el objeto de este trabajo, son:

1. Inicialmente, un menor conocimiento de las relaciones entre las variables que sí podemos encontrar en el MNL y, sobre todo en las RB. Esto es así por el concepto de “caja negra” inherente a las RN, al interpretar las relaciones de las capas ocultas del perceptrón multicapa con las entradas y salidas.
2. Se comprueba que no hay diferencias significativas en los resultados utilizando dos capas, y utilizando cuatro o seis, por lo que se opta por el modelo más sencillo con dos capas.
3. En el caso del error máximo absoluto, sigue habiendo valores muy próximos a 100 %, como es el caso de la conexión por ferrocarril con el puerto de Antofagasta, y las conexiones por carretera con los puertos de Arica, Ilo, Nuevapalmira, puerto Suárez, y Santos. Como se ha dicho en el caso anterior, estos porcentajes suponen un riesgo desde el punto de vista del cálculo de las elecciones modales, y, por tanto, las consecuencias derivadas para el proyecto en cuanto al cálculo de las captaciones del ferrocarril, ya que supone un mal dimensionamiento de las necesidades futuras de infraestructura. Eso implica desviaciones a su vez en los recursos necesarios para la operativa del servicio, para el cálculo de costes e ingresos, el coste total del proyecto y las necesidades o no de financiación.
4. En relación al error medio cuadrado, los valores son muy similares a las estimaciones del MNL, si bien, mejora en el caso de la conexión por carretera del

puerto de Arica, pasado de un 73,71% antes, a un 30,69% en las RN. Este indicador también mejora para las conexiones ferroviarias del puerto de Arica, Iquique y Santos, alrededor de un 19%, 5% y 3,5% respectivamente.

4.11. Calibración de la RB

A partir del mismo conjunto de datos usado para los modelos MNL en R, empleando las siguientes variables de red para cada par origen-destino.

1. Tiempos por alternativa
2. Coste por alternativa
3. Flujo anual por tipo de mercancía en toneladas métricas (MTPA)
4. Precio FOB (Free On Board) por producto o tipo de mercancía.
5. % de reparto por alternativa

Una diferencia con respecto a la codificación empleada para el ajuste de los modelos MNL es que en ellos se desarrolló un modelo independiente para cada tipo de mercancía, mientras que para la red bayesiana se empleó una sola red para todas, aunque se incluyó el factor precio FOB como indicador del tipo de mercancía (los graneles tienen bajos precios FOB y los contenedores valores altos).

La codificación de las variables se explica en la Tabla 4.6.

El proceso de calibración siguió los siguientes pasos, cuya base teórica está justificada y explicada en el apartado 2 de esta memoria:

1. Aplicación del test de Doornick-Hansen (DH) de normalidad, implementado en el paquete *asbio* de R, el cual mostró que las variables no se ajustaban adecuadamente a una distribución Normal.
2. Transformación de los datos mediante la transformación “Nonparanormal” propuesta por Lui et al. (2009), utilizando el paquete *huge* de R.
3. Aplicación de nuevo del test de Doornick-Hansen a la muestra transformada para verificar que efectivamente la transformación había proporcionado una muestra con distribución Normal.

Cuadro 4.6: Variables para la Red Bayesiana

Entrada	Variable	Descripción
Tiempo	$T_k = t_k - \min_k t_k$	Diferencia de tiempo con respecto al tiempo de la mejor alternativa k .
Coste	$C_k = c_k - \min_k c_k$	Diferencia de coste con respecto al coste de la mejor alternativa k .
Flujo de mercancía por año	$F_{i,j,m}$	Flujo de mercancía del origen i al destino j para la mercancía m .
Precio FOB	$FOB_{i,m}$	Precio FOB del producto en el punto de origen
Proporción de uso de la alternativa	$p_{m,i,j,k}$	Proporción de mercancía del tipo considerado entre el origen i y el destino j que hacen uso de la alternativa considerada k .

4. Como restricción para la calibración de la red bayesiana se incluyó que los nodos correspondientes a las proporciones $p_{m,i,j,k}$ no pudieran ser orígenes de los links del grafo. Esto se debe a que estas variables han de ser explicadas en todos los casos a partir del resto de variables de la red.
5. Aplicación del algoritmo max–min hill–climbing (mmhc) de Tsamardinos et al. (2016), para la calibración de la red bayesiana. Se aplica una lista de restricciones para que ningún arco de la red parta de los nodos correspondientes a los porcentajes de reparto en la elección (E_1 a E_{12}).
6. Aplicación de la función *arc.strength* en R, que permite conocer la importancia o fuerza de cada enlace de la red ajustada, sobre la red bayesiana resultante.
7. Obtención de los valores estimados de la red para las densidades condicionadas con la función *bn.fit* implementada en el paquete *bnlearn* de R.

El resultado de la calibración de la RB proporcionó los resultados de la Tabla 4.7.

Cuadro 4.7: Resultados de la calibración de la RB

Nodes: 42 Arcs: 61 Undirected arcs: 0 Directed arcs: 61 Average Markov blanket size: 4,86 Average neighbourhood size: 2,90 Average branching factor: 1,45 Learning algorithm: Max-Min Hill-Climbing	Constraint-based method: Max-Min Parent Children Conditional indep. test: Pearson's Correlation Score-based method: Hill-Climbing Score: BIC (Gauss.) Alpha threshold: 0,05 Penalization coefficient: 3,061246 Tests used in the learning procedure: 4504 Optimized: TRUE
--	--

Cuadro 4.8: Resultados ajuste RB

	Error Medio Cuadrado (RB)	Error Máximo Absoluto (RB)
ANTOFAGASTA_Road (carretera)	8,50 %	37,04 %
ANTOFAGASTA_Rail (ff.cc.)	7,40 %	39,53 %
ARICA_Road (carretera)	29,48 %	57,00 %
ARICA_Rail (ff.cc.)	4,05 %	48,81 %
ILO_Road (carretera)	12,26 %	46,24 %
IQUIQUE_Road (carretera)	13,27 %	39,66 %
IQUIQUE_Rail (ff.cc.)	2,20 %	46,96 %
NUEVAPALMIRA_Road (carretera)	15,83 %	47,30 %
PTOSUAREZ_Road (carretera)	3,55 %	46,86 %
PTOSUAREZ_Rail (ff.cc.)	2,20 %	46,96 %
SANTOS_Road (carretera)	10,27 %	48,10 %
SANTOS_Rail (ff.cc.)	6,30 %	50,15 %

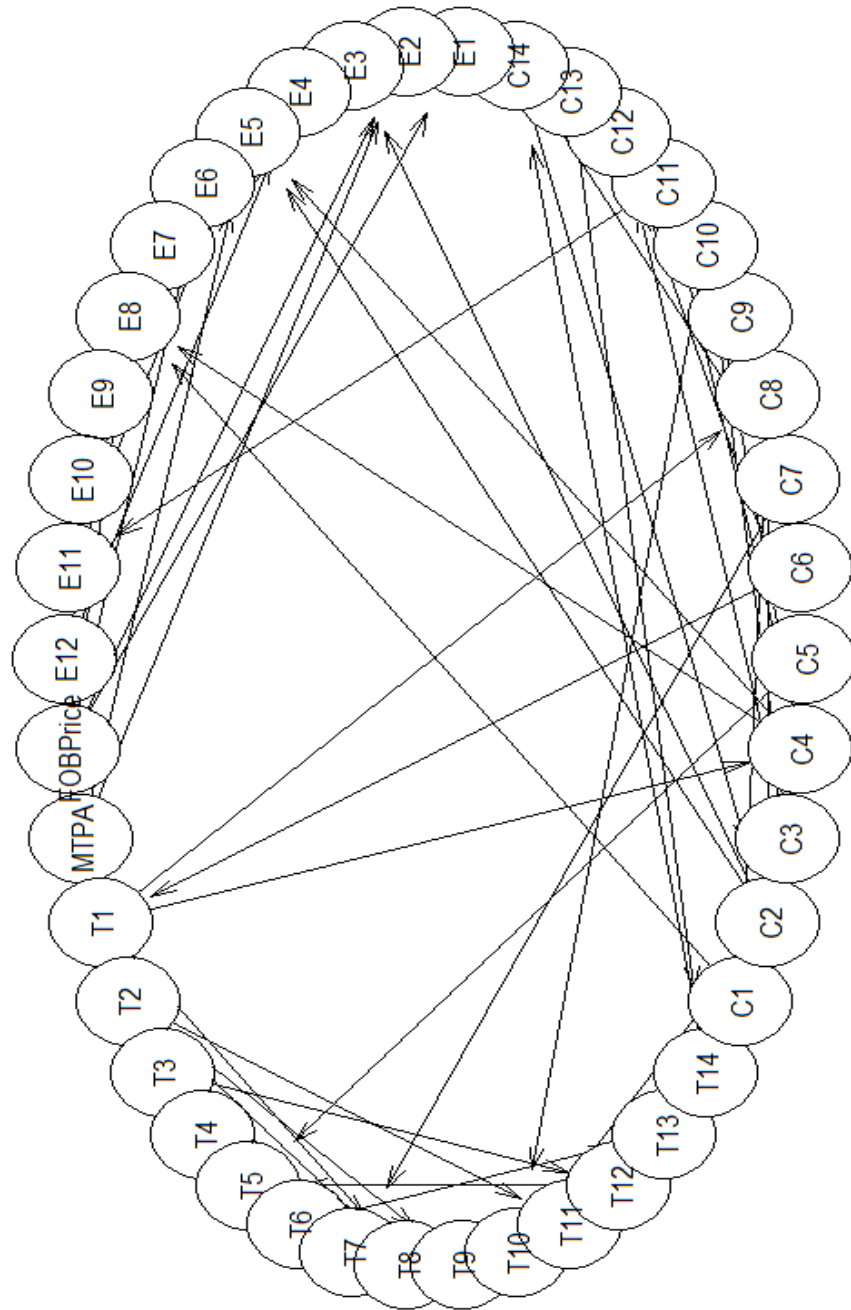


Figura 4.12: Grafo de la Red Bayesiana

Por otro lado, en la Figura 4.12 se muestra el grafo de la RB generada, donde se observan las relaciones entre los nodos de costes ($C_1 \dots C_{14}$) y tiempos ($T_1 \dots T_{14}$) para las distintas alternativas ($E_1 \dots E_{14}$), y cómo unos pocos de ellos en conjunto con los precios FOB y el flujo de mercancía por año (MTPA), explican la mayor parte de porcentajes de elección observados.

La figura de la RB obtenida nos da la relación entre las diferentes variables, por lo que su interpretación aporta más información que en el caso de las RN y MNL. Así, en el caso de las alternativas E1, E2, E3, E4, E6, E7, E10 y E12 las variables fundamentales son el precio FOB y la cantidad de toneladas transportadas (MTPA). Mientras que, en otras alternativas, E5, E8 y E11 (modo únicamente carretero), predomina el Coste asociado, bien de la propia alternativa, bien de un tramo común de otra en competencia. En el caso de los Tiempos de transporte, éstos dependen de tiempos de tramos igualmente en competencia, o de subtramos de alternativas mayores para el mismo modo de transporte (ferrocarril o carretera). Los Tiempos, a su vez, influyen en los costes totales de transporte, teniendo en cuenta la transformada monetaria de los mismos dependiendo de la utilización de un determinado número de recursos. Y en algunos casos, los costes derivan en un mayor o menor tiempo en el empleo de una u otra alternativa. Como ejemplo, el tiempo de la alternativa T1 (Arica carretera) influye en los costes de las alternativas C4 (Ilo carretera) y C8 (Arica ferroviaria), condicionado por la conexión existente en estos puertos por carretera y ferrocarril; y por otro lado, el coste de la alternativa C6 (Santos carretera), influye directamente en el tiempo de transporte T1 (Arica carretera), evidentemente, al ser el tramo de transporte más extenso de toda la red, entre la costa del Atlántico y la costa del Pacífico, habida cuenta de la desconexión ferroviaria actual entre la red andina y la red oriental en el tramo boliviano.

En la Tabla 4.8 se muestran los errores obtenidos de la estimación con la RB, para cada una de las alternativas analizadas:

Las principales conclusiones con este ajuste son, a efectos de error respecto a la elección modal de los datos observados reales, considerando que no se han utilizado estos datos para realizar ninguna asignación a red ya que no es el objeto de este trabajo, son:

1. Respecto al error cuadrático medio, todos los resultados mejoran las estimacio-

nes del MNL y la RN, salvo el caso de la conexión ferroviaria con los puertos de Arica, Iquique y Santos, que varían levemente respecto a la RN un 4 %, un 2 % y un 6 % respectivamente. Y en cuanto a las conexiones por carretera con los puertos de Nuevapalmira y Santos, igualmente suben respecto a la RN un 1 % y un 2 %, respectivamente. En general, errores muy pequeños.

2. Respecto al error absoluto máximo, de las doce alternativas, se mejora en nueve respecto de la RN, sólo empeorando en el caso de las conexiones ferroviarias de Arica, Iquique y Santos, entre un 40 % y un 50 % respecto de la RN. En el caso de Iquique, es un puerto con una demanda residual, por lo que su efecto es prácticamente nulo. En el caso de Santos (Brasil), es un tráfico de muy largo recorrido, unos 3.000 Km, por lo que es factible una desviación así en las captaciones. Por último, en el caso del puerto de Arica, probablemente la desviación venga producida por la variable de “preferencia” que incentiva esta conexión al ser el puerto que mejor operativa tiene de todos, por lo que la relación entre esta variables y las de tiempo y coste han influido en la estimación.
3. Queda demostrado que la RB arroja unas desviaciones, considerando los indicadores de bondad explicados, mejores que el MNL y la RN.

4.12. Ajuste en RStudio de RN y RB

La Figura 4.13 muestra el interfaz de RStudio donde ya se puede realizar el ajuste de las redes. Todo el código está implementado en el script *trainnet.R*. Basta con abrir este script y ejecutar los trozos de código deseados para ir obteniendo los resultados por pasos.

4.12.1. Ajuste de redes neuronales

El script para ajustar las redes neuronales usa la librería *neuralnet*.

El script está ordenado según los siguientes pasos:

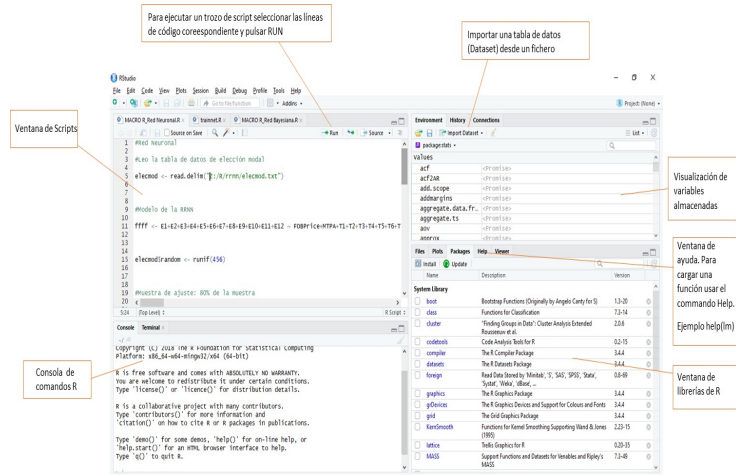


Figura 4.13: Interfaz RStudio

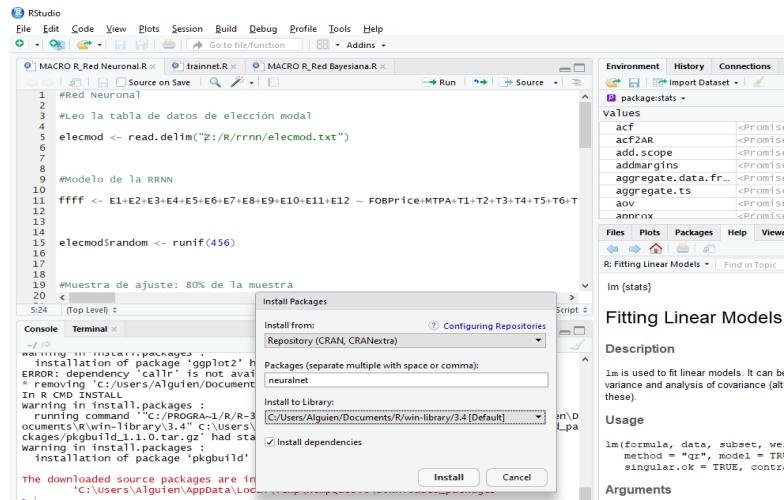


Figura 4.14: Instalación librería *neuralnet*

1. lee el fichero con los datos para la elección modal, los guarda en el data frame “*elecmod*” y genera el modelo para ajuste de la RN.

```
library(neuralnet)
#Leo la tabla de datos de elección modal
elecmod <- read.delim(C:/R/rrnn/elecmod.txt")
#Modelo de la RRNN
fff <- E1+E2+E3+E4+E5+E6+E7+E8+E9+E10+E11+E12
~ FOBPrice+MTPA+T1+T2+T3+T4+T5+T6+T7+T8+T9+T10+
T11+T12+T13+T14+C1+C2+C3+C4+C5+C6+C7+C8+C9+C10+C11+C12+C13+C14
computeSet <- data.frame(FOBPrice = elecmod$FOBPrice, MTPA = elecmod$MTPA,
T1 = elecmod$T1, T2 = elecmod$T2, T3 = elecmod$T3, T4 = elecmod$T4,
T5 = elecmod$T5, T6 = elecmod$T6, T7 = elecmod$T7, T8 = elecmod$T8,
T9 = elecmod$T9, T10 = elecmod$T10, T11 = elecmod$T11, T12 = elecmod$T12,
T13 = elecmod$T13, T14 = elecmod$T14, C1 = elecmod$C1, C2= elecmod$C2,
C3 = elecmod$C3, C4 = elecmod$C4, C5 = elecmod$C5, C6 = elecmod$C6,
C7 = elecmod$C7, C8 = elecmod$C8, C9 = elecmod$C9, C10 = elecmod$C10,
C11 = elecmod$C11, C12 = elecmod$C12, C13 = elecmod$C13, C14 = elecmod$C14,
E1 = elecmod$E1, E2 = elecmod$E2, E3 = elecmod$E3, E4 = elecmod$E4,
E5 = elecmod$E5, E6 = elecmod$E6, E7 = elecmod$E7, E8 = elecmod$E8,
E9 = elecmod$E9, E10 = elecmod$E10, E11 = elecmod$E11, E12 = elecmod$E12)
```

2. Genera la muestra para **ajuste de la red neuronal**, tal que el conjunto de datos se divide en dos muestras: una para el ajuste y otra para validación cruzada. El 80 % de los casos se asignan aleatoriamente al conjunto de datos de entrenamiento.

```
#Muestra de ajuste: 80 % de la muestra
computeSet$random <- runif
trainset <- subset(computeSet, random <= 0.8)
crossValSet <- subset(computeSet, random > 0.8)
computeSet2 <- computeSet[, 1:30]
crossValSet2 <- crossValSet[, 1:30]
```

3. Ajuste de las redes usando modelos con 2, 4, y 6 neuronas en la capa oculta. Para cada ajuste se guardan los resultados en los ficheros **netSol2.txt** y **sucesivos**. A) Usando el conjunto de datos **crossValSet2** entonces se guardan los resultados del ajuste solo para el conjunto de datos de validación. B) Usando **computeSet2** se guarda el conjunto de datos completo. El ajuste se hace por

el algoritmo de retropropagación (backpropagation) de los errores, como ya se indicó en la explicación teórica.

```
#Ajuste de los modelos de redes
nn2 <- neuralnet(ffff, trainset, hidden= c(2), rep=1, lifesign = "full", stepmax=1e6, threshold = 0.01)
rr <- compute(nn2, crossValSet2)
write.table(nn2$response, C:\R\rrnn\netSol2.txt")
nn4 <- neuralnet(ffff, trainset, hidden= c(4), rep=1, lifesign = "full", stepmax=1e6, threshold = 0.01)
rr <- compute(nn4, computeSet2)
write.table(netSol$response, C:\R\rrnn\netSol4.txt")
nn6 <- neuralnet(ffff, trainset, hidden= c(6), rep=1, lifesign = "full", stepmax=1e6, threshold = 0.01)
rr <- compute(nn6, computeSet2)
write.table(netSol$response, C:\R\rrnn\netSol6.txt")
```

4.12.2. Ajuste de redes bayesianas

El script para ajustar las redes neuronales usa la librería **bnlearn**.

1. Se aplica el test DH de normalidad y su transformación para normalizar los datos según se ha descrito, en caso de ser necesario.

```
#Redes Bayesianas
library(bnlearn)
#Test previos y normalización de la muestra
dh <- DH.test(computeSet, Y.names=names(computeSet))
tt <- DH.test(computeSet)
C <- huge.npn(computeSet)
C <- as.data.frame(C)
tt <- DH.test(C)
tt
```

2. Se ajusta la red bayesiana usando el dataframe “*blckl*” como lista de restricciones para el ajuste, tal que todos los arcos descritos por ese listado se eliminan del ajuste. Corresponde a los arcos con origen en los nodos E_i (con $i=1, \dots, 12$ alternativas de transporte), que corresponde a los porcentajes de elección con destino en el resto de variables. Así, la relación de causalidad se da en el sentido de que explicación del porcentaje de elección por el resto de las variables.

```
#Ajuste de la red bayesiana
```

```

rbg <- mmhc(C, blacklist = blackl)
bns <- arc.strength(rbg, C)
bns <- boot.strength(C)
bnf <- bn.fit(rbg, C)
bnstruct <- bn.net(bnf)

```

- Finalmente, se utiliza la RB para ajustar los datos de la muestra: se obtiene una predicción para cada variable de salida; se almacenan todas en el mismo data frame; por último son exportados.

```

#Generacion de la muestra con las predicciones
pp1 <- predict(bnf, .E1", C)
pp <- as.data.frame(pp1)
pp$pp1 <- pp$pp1 - C$E1
pp2 <- predict(bnf, .E2", computeSet)
pp$E2 <- pp2 - C$E2
pp3 <- predict(bnf, .E3", computeSet)
pp$E3 <- pp3 - C$E3
pp4 <- predict(bnf, .E4", computeSet)
pp$E4 <- pp4 - C$E4
pp$E5 <- predict(bnf, .E5", computeSet) - C$E5
pp$E6 <- predict(bnf, .E6", computeSet) - C$E6
pp$E7 <- predict(bnf, .E7", computeSet) - C$E7
pp$E8 <- predict(bnf, .E8", computeSet) - C$E8
pp$E9 <- predict(bnf, .E9", computeSet) - C$E9
pp$E10 <- predict(bnf, .E10", computeSet) - C$E10
pp$E11 <- predict(bnf, .E11", computeSet) - C$E11
pp$E12 <- predict(bnf, .E12", computeSet) - C$E12
predict(bnf, .E1", computeSet)
write.table(pp, "C:\\R\\rrnn\\rby.txt")

```

4.13. Resultados

Los resultados de Tabla 4.9 muestran el error cuadrático medio y el error máximo absoluto de cada alternativa de transporte, para cada método de estimación utilizado.

Las principales conclusiones con este ajuste son, a efectos de error respecto a la elección modal de los datos observados reales, considerando que en el caso de la RN y RB no se han utilizado estos datos para realizar ninguna asignación a red ya que no es el objeto de este trabajo, son:

Cuadro 4.9: Resultados comparativos ajuste MNL, RN y RB

	Error Medio Cuadrado		Error Máximo Absoluto	
	MNLogit	RN	MNLogit	RN
ANTOFAGASTA_Road (carretera)	15,08 %	14,389 %	99,48 %	92,214 %
ANTOFAGASTA_Rail (ff.cc.)	16,26 %	16,262 %	93,73 %	98,735 %
ARICA_Road (carretera)	73,71 %	30,690 %	100,00 %	97,492 %
ARICA_Rail (ff.cc.)	19,67 %	0,034 %	41,66 %	0,730 %
ILO_Road (carretera)	14,05 %	13,723 %	100,00 %	96,980 %
IQUIQUE_Road (carretera)	27,10 %	24,377 %	100,00 %	91,559 %
IQUIQUE_Rail (ff.cc.)	5,64 %	0,004 %	10,24 %	0,091 %
NUEVAPALMIRA_Road (carretera)	19,56 %	14,497 %	100,00 %	100,000 %
PTOSUAREZ_Road (carretera)	6,84 %	5,254 %	95,49 %	99,959 %
PTOSUAREZ_Rail (ff.cc.)	5,03 %	4,201 %	90,26 %	88,618 %
SANTOS_Road (carretera)	8,28 %	8,092 %	100,00 %	99,786 %
SANTOS_Rail (ff.cc.)	3,78 %	0,010 %	10,24 %	0,178 %
MEDIA	17,92 %	10,96 %	78,42 %	72,02 %
				46,22 %

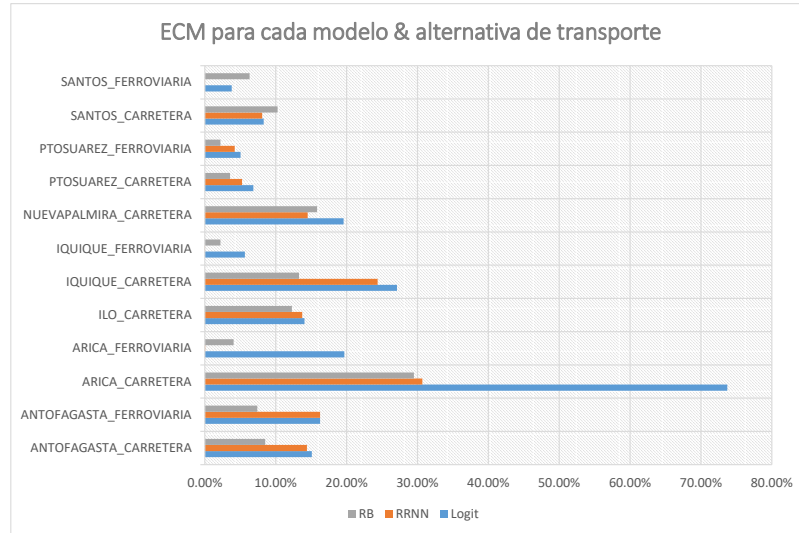


Figura 4.15: Error cuadrático medio por alternativa de transporte

1. Se observa una precisión similar para los modelos de RN y RB, si bien la RB mejora ligeramente esta última a la RN en las alternativas más utilizadas y con mayor flujo de mercancías: Arica por carretera, Antofagasta y Puerto Suárez.
2. El MNL es, en general, el modelo con menor precisión, aunque ofrece la ventaja de que sus parámetros son fáciles de interpretar y comprender.
3. A nivel global se aprecia que, en valor medio, la RB supera al MNL y a la RN. Ligeramente en el caso del error medio cuadrado de la RB respecto la RN, un 9,62 % contra un 10,96 %, y en mayor medida contra el MNL, un 17,92 %. Y de forma contundente en el caso del error máximo absoluto, un 46,22 % (RB) frente a un 72,02 % (RN), y un 78,42 % (MNL).
4. De forma gráfica, las Figuras 4.15 y 4.16 muestran lo observado en la tabla anterior, la RB supera a los otros métodos en la mayoría de los casos, especialmente en los más relevantes que son los indicados anteriormente: Arica por carretera, Antofagasta y Puerto Suárez.
5. Queda demostrado así que la aproximación bayesiana en el ajuste de las esti-

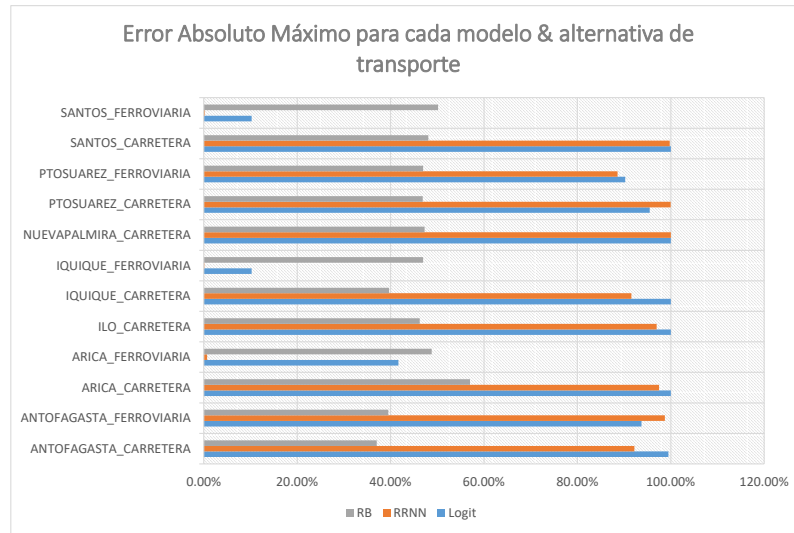


Figura 4.16: Error máximo absoluto por alternativa de transporte

maciones de la elección modal, para las alternativas planteadas en este modelo, es mejor que en el caso de las técnicas clásicas del MNL y algo más recientes de las RN.

Capítulo 5

Conclusiones

De los resultados obtenidos se pueden obtener las siguientes conclusiones.

1. En cuanto a la consecución de los objetivos generales y específicos definidos inicialmente:
 - a) *Objetivos generales:* el objetivo a.1 (mejorar el proceso de decisión de viajeros y operadores de carga) se alcanza ya que en este tipo de proyectos de planificación a futuro es habitual que exista una fase de participación pública y declaración de preferencias, en la que las entidades participantes (públicas y privadas) muestran y explican a los usuarios y operadores de carga las distintas alternativas de transporte. La aplicación de las redes bayesianas para la obtención de resultados de cada alternativa, en comparación con las otras técnicas, demuestra en este caso que ofrece mejores ajustes en las variables principales de tiempo y coste por alternativa. Por tanto, en este sentido, mejora el proceso de decisión de viajeros y operadores de carga, a partir del análisis detallado de la información de partida, en dicho proceso de participación pública. Esto repercute posteriormente en el análisis multicriterio Analytic Hierarchy Process (AHP) de las alternativas, a partir de los factores principales de cada una, con sus ponderaciones correspondientes aplicadas en el proceso de planificación. En el caso de estudio mostrado en concreto, a partir de las ton/año de cada alternativa se determinan los porcentajes de reparto observados reales

entre alternativas. Este valor es fundamental ya que determina las probabilidades de uso reales por sentido import/export, tipo de mercancía y par OD, para cada una de las combinaciones ID1 e ID2. El resultado es el porcentaje obtenido de la división de toneladas/año para cada combinación de las anteriormente indicadas, entre la suma total de carga por OD. En base a estos porcentajes, posteriormente se determinan los errores relativos y absolutos del Multi Nomial Logit, Red Neuronal y Red Bayesiana, lo que muestra la bondad en el ajuste de cada una de las técnicas utilizadas. Por tanto, a menor error obtenido, mejor ajuste para cada alternativa de transporte elegida por los operadores de carga, lo cual también se puede extrapolar a la decisión de los viajeros. Así, la minimización de los errores se puede transformar en unidades monetarias: 1) a partir del coste de cada alternativa de transporte, por lo que un mejor ajuste en los errores implica una reducción en los costes individuales por alternativa, y por tanto, en el sumatorio del coste total del proyecto; 2) un mejor ajuste permite saber exactamente los ingresos obtenidos de cada elección de alternativa. Por tanto, se consigue un mejor ajuste en el ratio coste-beneficio; en cuanto a los objetivos a.2 (optimizar los modelos de transporte utilizados en planificación) y a.3 (comprobar que la aplicación de las aproximaciones bayesianas en la etapa de reparto modal supone obtener mejores resultados), si bien se han comprobado en este estudio, no se han podido comprobar con otros proyectos de esta envergadura (bioceánica) y corredores multimodales de transporte de mercancía a nivel mundial. La aplicación posterior de esta aproximación en proyectos similares o más pequeños permitirá comprobar la asunción de generalidad o no. El caso estudiado en el presente trabajo supone un primer paso para ello; el objetivo a.4 (promover las redes bayesianas en entornos complejos que ya son una realidad como son los Sistemas Inteligentes de Transporte (ITS), los Sistemas Cooperativos (V2V, V2I, V2I2V, I2I) de transporte e infraestructuras, y los vehículos autónomos) podría alcanzarse si las publicaciones derivadas de la tesis tienen impacto en la comunidad interesada, entidades públicas y privadas, y en sus proyectos de planificación de in-

fraestructuras. Ese el motivo para continuar y avanzar con esta línea de investigación a futuro.

- b) **Objetivos específicos:** En cuanto a los objetivos b.1 (conseguir mejores resultados con menor cantidad de información), b.2 (disminuir los tiempos de procesamiento para el reparto modal) y b.4 (mejorar los algoritmos utilizados habitualmente), se consiguen teniendo en cuenta tres de los principales factores de la base matemática de las redes bayesianas: la relación entre los nodos de las variables padres y de las variables hijos, la estructura de dependencia probabilística entre ellas, y la posibilidad en cuanto a la utilización de información cualitativa y cuantitativa, hacen que los pesos de las variables así como los ajustes iterativos se optimicen con la información disponible a lo largo del tiempo en comparación con los otros métodos. Además de esto, los resultados obtenidos mediante la aproximación bayesiana mejoran los obtenidos mediante los algoritmos habituales; en cuanto al objetivo b.3 (mejorar el ajuste coste-beneficio derivado de las mejoras aportadas por las RB en los resultados, cara a optimizar los recursos económicos existentes y conseguir financiación para los proyectos de infraestructura y transporte), también se consigue, ya que si bien el detalle de los costes de cada alternativa de transporte está sujeto a confidencialidad, minimizar los errores de elección de cada alternativa supone ajustar los costes y los beneficios derivados del uso de cada una, por lo que se mejora el ajuste de ambas cifras. La ficha descriptiva del proyecto correspondiente al COSIPLAN (Consejo Suramericano de Infraestructura y Planeamiento) de UNASUR (Unión de Naciones Suramericanas): http://www.iirsa.org/proyectos/detalle_proyecto.aspx?h=1351, permite comprobar que es uno de los proyectos más importantes del IIRSA (Iniciativa para la Integración de la Infraestructura Regional Suramericana) con una inversión total de unos 7.000 millones de USD.

2. En cuanto a los resultados obtenidos, a nivel general, por el momento no permiten una generalización para modelos clásicos de cuatro etapas, de transporte de mercancías. Son específicos para este caso concreto, en el que se ha utilizado una RB en la etapa de reparto modal. Teniendo en cuenta, además, que la

confidencialidad del uso de los mismos condiciona el detalle de las alternativas propuestas y por tanto, de los resultados mostrados. Al respecto, mencionar que en el capítulo 4.3 se enumeran otros trabajos similares recientes sobre modelos de transporte como los de Sun et al. (2006), Correa et al. (2009), Tang et al. (2012), aplicando RB a modelos de transporte de vehículo privado y pasajeros, Daziano et al. (2013), donde están contenidas las RB asociadas a la distribución multinomial, Yannis Tyrinopoulos (2013) y Tai Yu Ma (2015) donde pueden verse distribuciones más generales aplicadas a este tipo de redes, con resultados igualmente prometedores. Sin embargo, todos ellos se focalizan en ámbitos de estudio urbanos o metropolitanos y de vehículo privado, bien para analizar temas como: estimaciones de viajes de vehículos privados (pasajeros); sistemas de peaje en autopistas; la seguridad vial, considerando los beneficios derivados del uso de modelos Markov Chain Monte Carlo (MCMC) y de los modelos bayesianos Empirical Bayes (EB) y Full Hierarchical Bayes (FHB); la movilidad en situaciones de congestión vehicular, con modelos Probit; o los patrones de movilidad urbana de los trabajadores (commuters y similares) de zonas frontera metropolitanas, analizando cuáles son las principales variables que influyen en un modelo de reparto modal orientado a la promoción del transporte público en detrimento del vehículo privado. Por tanto, bajo esta consideración, si bien son similares en cuanto a la orientación de la elección modal, difieren en cuanto a su uso hacia un modelo de transporte de mercancías como el indicado en el caso de estudio que se presenta, el cual además es un caso real.

3. La principal contribución original de este trabajo es el uso de una RB en la etapa de reparto modal de un modelo clásico de cuatro etapas, aplicado a un modelo multimodal de mercancías de ámbito mundial.
4. Por lo tanto, la realización de este trabajo permite afirmar la hipótesis inicial de partida: “las Redes Bayesianas son una mejor aproximación en los modelos de reparto modal, incluidos a su vez de modelos de transporte, que los habitualmente utilizados como los modelos Logit y Multinomial Logit, e incluso que las Redes Neuronales”.
5. En el caso concreto de un modelo de transporte de mercancías de ámbito

mundial se ha verificado pese a la falta de evidencias iniciales y los resultados obtenidos han servido para apoyar dicho planteamiento.

6. Se demuestra la idea de que no todas las variables de una RB influyen de igual manera en la variable de interés. Es importante recalcar en este contexto, que el objetivo del procedimiento consiste en identificar entre las variables disponibles, cuál o cuáles, intervienen más para disminuir la entropía del objetivo, entendiendo esto como una disminución de la incertidumbre.
7. Los modelos de transporte son una poderosa herramienta para ayudar en la toma de decisiones relacionadas con las infraestructuras y las inversiones que un país ha de llevar a cabo a futuro, por lo que mejorar las estimaciones en este tipo de proyectos de infraestructura supone.
8. La utilización conjunta de algoritmos y métodos matemáticos con herramientas y tecnologías de modelización, aplicadas a la planificación de transportes a partir de la utilización de datos históricos o en tiempo real, es una realidad que se ha de imponer en los próximos años para la optimización de los recursos e infraestructuras, mejora del medio ambiente, del ahorro energético, reducción de la huella de carbono, lo que en general significa mejorar la calidad de vida de la población.
9. Tanto la definición y elección de los datos necesarios, como de la metodología a aplicar son aspectos fundamentales que condicionan la información disponible y los posibles resultados a obtener. Éstos se tienen que ajustar a los objetivos planteados, de cara a conseguir la mayor precisión posible, siendo a su vez lo más realistas posibles.
10. Una optimización en la planificación, utilizando aproximaciones matemáticas novedosas, como es en este caso la aplicación de Redes Bayesianas, no sólo significa mejorar otros procedimientos si no avanzar en nuevas líneas que son igual de consistentes desde el punto de vista teórico.
11. La reducción de los errores en los ajustes implica directamente una mejora en los recursos económicos de un país o entidad pública, de red o de servicio,

que se establezcan en cada escenario. Es decir, existe una transformada de los datos empíricos a los aspectos económicos, lo que significa conseguir una monetización de las mejoras obtenidas.

12. Por lo tanto, los modelos matemáticos son fundamentales en la toma de decisiones. En este caso, para calcular los posibles niveles de endeudamiento, financiación exterior, establecimiento de las tarifas, tipos de vehículo, condiciones de servicio, condiciones administrativas, necesidad de cambios normativos y legislativos que den soporte a los gobiernos, empresas y entidades, cambios en las políticas sociales, etc.
13. La construcción de modelos configurables de transporte de pasajeros y mercancías implica la posibilidad de analizar y evaluar sistemas multimodales de transporte teniendo en cuenta un número considerable de factores y variables, las cuales pueden presentar relaciones directas o indirectas.
14. La aplicación de la metodología basada en un modelo de cuatro etapas permite, en el caso de estudio concreto de un posible corredor ferroviario entre Brasil y Perú (CFBC), obtener un modelo capaz de estimar las futuras captaciones de tráfico del ferrocarril respecto otros modos.
15. La herramienta desarrollada ha sido adoptada por la Unidad Técnica de Ferrocarriles (UTF) del Viceministerio de Transporte (VMT) como apoyo en la planificación de transporte, si bien, a día de hoy este proyecto se encuentra a la espera de cambios por la situación interna y mundial existente.
16. La implementación de un modelo de transporte en un software específico (en este caso TransCAD) proporciona las siguientes ventajas:
17. Contar con una base de datos GIS adecuadamente estructurada con la información relevante del modelo.
18. Automatizar el proceso de ejecución del modelo de tal forma que personal con una formación básica en modelos de transporte pueda hacer uso de este.
19. Agilizar el análisis de escenarios y el reajuste de las previsiones a medida que se cuente con nuevos datos.

20. Proporcionar información gráfica que facilite la interpretación de los resultados.
21. Se ha evidenciado el hecho teórico de que no todas las variables de una RB influyen de igual manera en la variable de interés, y en este caso concreto en las alternativas planteadas.
22. La aproximación bayesiana empleada permite comparar tanto las relaciones entre variables como el peso de cada una de ellas respecto de variables alternativas, ya sea para la red inicialmente construida, o para otras redes distintas, en base a los datos y bondad de la información disponible.
23. Los avances algorítmicos y de procesamiento de grandes cantidades de información están permitiendo que nuevas técnicas, como las Redes Bayesianas, se postulen como alternativas reconocidas, obteniéndose resultados prometedores.
24. La aplicación de diferentes tests, tanto el de Doornick-Hansen para aceptar o rechazar la hipótesis de Normalidad de la muestra de datos, como el transformación Nonparanormal en el caso de rechazo de este supuesto, suponen un avance en la fase de análisis y depuración de la información de partida. Esto implica no solo conocer las funciones de distribución asociadas, si no la posibilidad de transformación en el caso de no cumplir el criterio de Normalidad, sin pérdida de variabilidad, lo cual se traduce en la posibilidad de utilización de las RBG.
25. En el caso del estudio descrito, los tres modelos de elección discreta empleados (Logit Multinomial (MNL) – Red Neuronal (RN) – Red Bayesiana (RB)) permiten obtener estimaciones del reparto modal para el modelo de transporte construido. Lo cual implica cumplir un supuesto básico de partida.
26. Los resultados y ajustes obtenidos mediante la aplicación de Redes Bayesianas, en comparación con los modelos Logit Multinomiales y de Redes Neuronales, permite minimizar los errores a la hora de estimar las mejores alternativas de transporte entre los modelos de elección modal empleados.
27. En relación al MNL, son las siguientes:

- a) Principalmente, el valor del error medio cuadrado y el error máximo absoluto para la conexión con el puerto de Arica por carretera son muy elevados. Esto es peligroso ya que se ha demostrado que este puerto es una de las salidas más importantes de las exportaciones hacia el océano Pacífico. Por tanto, desviaciones así repercuten en un mal dimensionamiento de los recursos e infraestructuras, y supone una mala planificación a futuro y unos condicionantes económicos y financieros de inversión mucho más elevados a largo plazo.
- b) En relación a la otra alternativa principal de salida, el puerto de Antofagasta, el error medio cuadrado obtenido está en unos valores razonables, si bien, los errores máximos absolutos son muy elevado, entre un 93 % y un 99,5 %, lo cual es peligroso para las estimaciones futuras.
- c) Además, se suma el desconocimiento de las relaciones existentes de dependencia entre las variables significativas, que sí existen en las Redes Bayesianas.
- d) Las conclusiones de los anteriores gráficos a partir de las estimaciones del modelo obtenidas en la anterior tabla, tomando de ejemplo las exportaciones bolivianas por el océano Pacífico:
- e) Situación base: la vía principal de salida de la carga es la conexión hasta el puerto de Antofagasta y desde aquí por vía marítima. Hasta este puerto el tráfico llega en modo ferroviario. También se realizan salidas, pero menores, por Arica e Ilo, conectando por carretera.
- f) Situación futura: con el corredor ferroviario funcionando al 100 %, éste se posiciona como un modo competitivo frente a la carretera, en costes y tiempos, para el transporte de mercancías. De esta forma, las mercancías se reparten entre Antofagasta y el puerto de Ilo. En este último caso se aprecia como el flujo de carga es mayor por ferrocarril que por carretera, con lo que el total de carga transportada aumenta respecto la situación base.
- g) Por tanto, las funciones de utilidad resultantes de la aplicación del MNL para calcular la elección modal son consistentes en la situación base, y

por tanto, pueden ser usadas para las proyecciones de carga en situación futura. Esto es importante en la etapa de asignación a la hora de calcular las captaciones posibles en el corredor ferroviario.

28. En relación a la RN, considerando que no se han utilizado estos datos para realizar ninguna asignación a red ya que no es el objeto de este trabajo, son las siguientes:

- a) Inicialmente, un menor conocimiento de las relaciones entre las variables que sí podemos encontrar en el MNL y, sobre todo en las RB. Esto es así por en concepto de “caja negra” inherente a las RN, al interpretar las relaciones de las capas ocultas del perceptrón multicapa con las entradas y salidas.
- b) Se comprueba que no hay diferencias significativas en los resultados utilizando dos capas, y utilizando cuatro o seis, por lo que se opta por el modelo más sencillo con dos capas.
- c) En el caso del error máximo absoluto, sigue habiendo valores muy próximos a 100 %, como es el caso de la conexión por ferrocarril con el puerto de Antofagasta, y las conexiones por carretera con los puertos de Arica, Ilo, Nuevapalmira, puerto Suárez, y Santos. Como se ha dicho en el caso anterior, estos porcentajes suponen un riesgo desde el punto de vista del cálculo de las elecciones modales, y, por tanto, las consecuencias derivadas para el proyecto en cuanto al cálculo de las captaciones del ferrocarril, ya que supone un mal dimensionamiento de las necesidades futuras de infraestructura. Eso implica desviaciones a su vez en los recursos necesarios para la operativa del servicio, para el cálculo de costes e ingresos, el coste total del proyecto y las necesidades o no de financiación.
- d) En relación al error medio cuadrado, los valores son muy similares a las estimaciones del MNL, si bien, mejora en el caso de la conexión por carretera del puerto de Arica, pasado de un 73,71 % antes, a un 30,69 % en las RN. Este indicador también mejora para las conexiones ferroviarias del puerto de Arica, Iquique y Santos, alrededor de un 19 %, 5 % y 3,5 % respectivamente.

29. En relación a la RB, considerando que no se han utilizado estos datos para realizar ninguna asignación a red ya que no es el objeto de este trabajo, son las siguientes:
- a) Se observa una precisión similar para los modelos de RN y RB, si bien la RB mejora ligeramente esta última a la RN en las alternativas más utilizadas y con mayor flujo de mercancías: Arica por carretera, Antofagasta y Puerto Suárez.
 - b) El MNL es, en general, el modelo con menor precisión, aunque ofrece la ventaja de que sus parámetros son fáciles de interpretar y comprender.
 - c) A nivel global se aprecia que, en valor medio, la RB supera al MNL y a la RN. Ligeramente en el caso del error medio cuadrado de la RB respecto la RN, un 9,62 % contra un 10,96 %, y en mayor medida contra el MNL, un 17,92 %. Y de forma contundente en el caso del error máximo absoluto, un 46,22 % (RB) frente a un 72,02 % (RN), y un 78,42 % (MNL).
 - d) De forma gráfica, igualmente se observa que la RB supera a los otros métodos en la mayoría de los casos, especialmente en los más relevantes que son los indicados anteriormente: Arica por carretera, Antofagasta y Puerto Suárez.
 - e) En este sentido, la figura de la RB obtenida nos da la relación entre las diferentes variables, por lo que su interpretación aporta más información que en el caso de las RN y MNL. Así, en el caso de las alternativas E1, E2, E3, E4, E6, E7, E10 y E12 las variables fundamentales son el precio FOB y la cantidad de toneladas transportadas (MTPA). Mientras que, en otras alternativas, E5, E8 y E11 (modo únicamente carretero), predomina el Coste asociado, bien de la propia alternativa, bien de un tramo común de otra en competencia. En el caso de los Tiempos de transporte, éstos dependen de tiempos de tramos igualmente en competencia, o de subtramos de alternativas mayores para el mismo modo de transporte (ferrocarril o carretera). Los Tiempos, a su vez, influyen en los costes totales de transporte, teniendo en cuenta la transformada monetaria de los mismos dependiendo de la utilización de un determinado número de

recursos. Y en algunos casos, los costes derivan en un mayor o menor tiempo en el empleo de una u otra alternativa. Como ejemplo, el tiempo de la alternativa T1 (Arica carretera) influye en los costes de las alternativas C4 (Ilo carretera) y C8 (Arica ferroviaria), condicionado por la conexión existente en estos puertos por carretera y ferrocarril; y por otro lado, el coste de la alternativa C6 (Santos carretera), influye directamente en el tiempo de transporte T1 (Arica carretera), evidentemente, al ser el tramo de transporte más extenso de toda la red, entre la costa del Atlántico y la costa del Pacífico, habida cuenta de la desconexión ferroviaria actual entre la red andina y la red oriental en el tramo boliviano.

- f) La RB y el MNL permiten interpretar los parámetros resultantes. La RN en menor medida por el efecto “caja negra” mencionado en relación a las capas ocultas del modelo multicapa.
 - g) Queda demostrado así que la aproximación bayesiana en el ajuste de las estimaciones de la elección modal, para las alternativas planteadas en este modelo, es mejor y más precisa, que en el caso de las técnicas clásicas del MNL y algo más recientes de las RN, tanto en base al error cuadrático medio como a los errores máximos de ajuste.
30. En definitiva, se han respondido a las preguntas clave formuladas durante el desarrollo del trabajo:
- a) ¿Cuáles son las contribuciones originales del trabajo? El uso de las redes bayesianas en la etapa de reparto modal de un modelo de transportes (clásico de cuatro etapas) de mercancías, multimodal, a nivel mundial, confirmando que los resultados obtenidos son mejores que los obtenidos con otras técnicas más habituales, como los modelos multinomial logit (MNL) y las redes neuronales (RN).
 - b) ¿Cuál es su naturaleza? La definición de la temática y el planteamiento del problema de investigación centran el estudio en la planificación de las infraestructuras a futuro. En este caso concreto, de una infraestructura multimodal con relaciones a todos los niveles posibles: distritales, regionales, estatales y continentales. El desarrollo de la perspectiva teórica se

basa en las Redes Bayesianas, en la definición, tipología y aplicación de las mismas, considerando sus principales características y propiedades. La recolección, análisis e interpretación de los datos toma de referencia un caso práctico de estudio, un corredor ferroviario que una el océano Atlántico con el océano Pacífico, prioritariamente para el transporte de mercancías. Siendo las variables principales en la elección de las alternativas de transporte multimodal las de tiempo y coste, la naturaleza de la investigación es meramente cuantitativa. Si bien, la posibilidad de incluir otro tipo de variables cualitativas hace más extensa esta naturaleza inicial. La estrategia metodológica se basa en la comparativa de los resultados obtenidos para las distintas alternativas de transporte, en cuanto a la elección del modo de transporte (reparto modal), en el modelo de transportes construido a tal efecto. Las metas del estudio, finalmente, han sido las de comprobar que los resultados obtenidos con las redes bayesianas son mejores que las obtenidas con otros métodos habituales, los multinomial logit (MNL) y las redes neuronales (RN).

- c) ¿Cuál es su alcance? a) macroscópico, como es el caso de estudio presentado, pudiéndose aplicar de forma añadida tanto a la etapa (1) de generación-atracción, como a la etapa (2) de distribución de viajes (transformada de la matriz de generación-atracción a la matriz de origen-destino mediante una función de impedancia). El siguiente paso, lógico, de la presente investigación precisamente va en línea con la implementación de este tipo de redes bayesianas en tres de las cuatro etapas de un modelo de mercancías. b) microscópico, en los modelos de tráfico de ámbito más urbano o interurbano, pudiéndose aplicar a los modelos de elección de ruta o seguimiento vehicular, inicialmente.
- d) ¿Cuáles son sus limitaciones? La principal limitación es, hasta el momento, la inexistencia de proyectos similares de transporte de mercancías en el que se hayan podido replicar el uso de redes bayesianas al reparto modal, que verifiquen que los resultados obtenidos en este caso. Igualmente, la difusión de las redes bayesianas como herramientas de aplicación en los proyectos de planificación de infraestructura todavía tiene un margen de

penetración en las entidades públicas y empresas de ingeniería, por lo que a día de hoy su uso no está muy extendido.

31. Estos resultados abren una nueva línea de aplicación para las RB en los problemas de elección modal en transporte. Por un lado, ofrecen posibilidades interesantes en cuanto al establecimiento de relaciones de causalidad entre las variables del problema y por otro lado, los resultados en este caso en particular, muestran que las RB ofrecen un gran potencial para ser más precisas que los modelos clásicos de reparto modal que se han venido utilizando hasta la fecha.

Capítulo 6

Futuras líneas de investigación

A continuación, se muestran algunas de las principales líneas de investigación que pueden derivarse del presente trabajo con Redes Bayesianas:

1. Mejora de los modelos macroscópicos de 4 etapas en las otras etapas en fase de estudio, fundamentalmente la etapa 1-Generación-Atracción y la etapa 2-Distribución
2. Mejora de los algoritmos de seguimiento vehicular en los modelos microscópicos e híbridos de simulación de tráfico, partiendo de datos obtenidos en tiempo real de los centros de control.
3. Optimización en las técnicas matemáticas empleadas en los modelos de eventos discretos, en relación al tratamiento de las longitudes máxima de cola y los servidores de, mejorando los tiempos en los servicios. Un ejemplo claro son las terminales portuarias en las que la optimización del stockage es fundamental para garantizar los tiempos y recursos necesarios.
4. Sistemas Expertos aplicados al COVID19. Un ejemplo es el desarrollado en el grupo de trabajo /investigación MATGEN del CEMAT, dentro de la iniciativa de Acción Matemática contra el Coronavirus. Se está trabajando en un sistema experto con metodología bayesiana y big data para la detección de puntos de inflexión en series temporales aplicación a las series de datos del covid-19 (<http://matematicas.uclm.es/cemat/covid19/>). EOI sobre el SARS-COV-2 y la enfermedad COVID19 - (Unión Europea+Instituto de salud Carlos III).

5. Mejora de las técnicas aplicadas en Digital Humanities (interpretación de legajos manuscritos anteriores al siglo XIX) en el procesamiento de la información obtenida de imágenes de alta resolución.
6. Sistemas Inteligentes de Transporte (ITS) aplicados a Transporte Público y Privado y Sistemas Cooperativos vehículo-infraestructura (V2V, V2I, V2I2V, I2I), en cuanto a la mejora de los algoritmos utilizados entre ambos “objetos” en tiempo real. Supone un gran reto de cara a la implantación de los sistemas 5G de comunicación y la sensorización total de las ciudades y extensión de las Smart Cities.
7. Big Data y Soft Data aplicados al transporte multimodal e intermodal (viario, portuario, aéreo) de mercancías, de cara a una adecuada planificación de las infraestructuras, políticas tarifarias y de costes.
8. Modelos de Transporte con Vehículos de conducción Autónoma, RPAS (drones) y sistemas robóticos.
9. Redes sociales, redes de telecomunicación, redes de abastecimiento y distribución (electricidad, agua, gas. . .) en cuanto al ajuste de los modelos de consumo y de demanda.
10. Innovación en redes eléctricas de recarga urbana e interurbana de vehículos, tanto en entornos residenciales como comunitarias, bien por suministro directo en punto fijo como inducido de forma dinámica. Un ejemplo son los sistemas de recarga por oportunidad en cabeceras aplicados a sistemas BRT (Bus Rapid Transit) con pantógrafos.
11. Modelos genéticos, modelos tumorales, innovación médica. Mejora en los algoritmos de diagnóstico de imágenes tumorales, comparando las Redes Bayesianas con las Redes Neuronales estándar y convolucionales recientemente aplicadas en enfermedades degenerativas, como el Alzheimer.

Al respecto, entre los proyectos europeos actuales relacionados con la movilidad cabe destacar: **C-Roads Spain, Bamboo, Levitate y Momentum**, los cuales se están desarrollando o se van a desarrollar con el horizonte 2022.

1. (<https://intercor-project.eu/homepage/c-roads/>) C-ROADS: Plataforma Abierta para mejorar la conectividad europea de los ITS (Connecting Europe Facility (CEF)) C-ITS
2. BAMBOO: Big Data Analytics for Mobility Modelling
(<https://www.eurostars-eureka.eu/project/id/12063>,
<https://www.era-learn.eu/network-information/networks/eurostars-2/eurostars-cut-off-8/big-data-analytics-for-mobility-modelling>)
3. LEVITATE: Societal Level Impacts of Connected and Automated Vehicles (levitate-project.eu)
4. (<https://cordis.europa.eu/project/rcn/221856/factsheet/en>) MOMENTUM: Proyecto para modelizar soluciones emergentes de transporte en ciudades de la UE: tecnologías disruptivas, MaaS, CAVs (connected and autonomous vehicles).

A estos proyectos, también hay que añadir los que se están desarrollando en Reino Unido orientados a los vehículos autónomos y conectados (CAVs), entre los que destacan: **Flourish** y **Human Drive**.

1. FLOURISH: (<http://www.flourishmobility.com/>)
2. HUMAN DRIVE: (<https://humandrive.co.uk/>)

Capítulo 7

Bibliografía

1. Aho, K. (2014). *asbio: A Collection of Statistical Tools for Biologists. Contains functions from: Foundational and Applied Statistics for Biologists using R*, (<https://cran.r-project.org/web/packages/asbio/index.html>). CRC/Taylor and Francis, Boca Raton, FL
2. Aldrich, J.A., Nelson, F.D. and Adler, E.S. (1984.). *Linear Probability, Logit, and Probit Models*. Sage Publications, inc.
3. Anderson, T.W. (2003). *An introduction to multivariate statistical analysis (3rd ed.)*. Wiley-Interscience.
4. Bayes, T.; Price, Mr. (1764). *An Essay towards solving a Problem in the Doctrine of Chances*. Philosophical Transactions of the Royal Society of London 53: 370–418. doi:10.1098/rstl.1763.0053
5. Berge, C. (1973). *Graphs and Hypergraphs*. Amsterdam: North-Holland.
6. Bowman, K.O. and Shenton, L.R. (1975). *Omnibus test contours for departures from normality based on $\sqrt{b_1}$ and b_2* . Biometrika, vol 2, 103-117.
7. Bromley, J. (2005). *Guidelines for the use of Bayesian networks as a participatory tool for Water Resource Management*. Centre for Ecology and Hydrology, C01704, 117.

8. Castillo, E., Gutiérrez, J.M. y Hadi, A.S. (1997). *Expert Systems and Probabilistic Network Models*. Springer Verlag.
9. Castillo, E. y Kjærulff, U. (2003). *Sensitivity analysis in Gaussian Bayesian networks using a symbolic-numerical technique*. Reliability Engineering and System Safety, 79, 139-148.
10. Clancey, W.J. (1993). *Notes on Heuristic classification*. Artificial Intelligence, 59: 191-196.
11. Cobb, B.R., Rumí, R. and Salmerón, A. (2007). *Bayesian networks models with discrete and continuous variables*. *Advances in Probabilistic Graphical Models, Studies in Fuzziness and Soft Computing*. Springer 81-102.
12. Correa, M., Bielza, C., Pamies-Teixeira, J. (2009). *Comparison of Bayesian networks and artificial neural networks for quality detection in a machining process*. Expert Systems with Applications, Volume 36, Issue 3, Part 2, Pages 7270-7279, ISSN 0957-4174. DOI.org/10.1016/j.eswa.2008.09.024.
13. Cowell, R. G., Dawid, A. P., Lauritzen, S. L. y Spiegelhalter, D. J. (1999). *Probabilistic Networks and Expert Systems*. Springer, Barcelona.
14. Cowell, R. G. (2005). *Local Propagation in Conditional Gaussian Bayesian Networks*. Journal of Machine Learning Research, 6, 1517-1550.
15. Daly, A., Zachary, S., Hensher, D. and Dalvi, Q. (1978). *Improved multiple choice models*.
16. D'Ágostino, R. B. (1970). *Transformation to normality of the null distribution of g_1* . Biometrika, 57, 679-681.
17. Daziano, R., Miranda-Moreno, L. & Shahram Heydari (2013). *Computational Bayesian Statistics in Transportation Modelling: From road Safety Analysis to Discrete Choice*, Transport Reviews, Vol. 33.
18. Dean, T. y Wellman, M. (1991). *Planning and Control*. San Mateo, California: Morgan Kaufmann.

19. Doornik, J.A. and Hansen, H. (2008). *An Omnibus Test for Univariate and Multivariate Normality*. Oxford Bulletin of Economics and Statistics, 70, 927-939
20. Golumbic, M. C. (1980). *Algorithmic Graph Theory and Perfect Graphs*. London: Academic Press.
21. Gómez-Villegas, M.A., Maín, P. y Susi, R. (2006). *Sensitivity analysis of extreme inaccuracies in Gaussian Bayesian Networks*. Proceedings of the Third European Workshop on Probabilistic Graphical Models, Prague, Czech Republic, pp. 139-146.
22. Gomez-Villegas, M.A., Main, P. and Susi, R. (2007). *Sensitivity analysis in Gaussian Bayesian networks using a divergence measure*. Communications in Statistics- Theory and Methods, vol 36, 523-539.
23. Gómez-Villegas, M.A., Maín, P. y Viviani, P. (2014). *Sensitivity to evidence in Gaussian Bayesian networks using mutual information*. Information sciences 275, 115-126
24. Gómez-Villegas, M.A., González- Pérez, B., De Gregorio Vicente, O (2019). *Redes bayesianas aplicadas a la nueva movilidad mundial*. XXXVIII Congreso Nacional de Estadística e Investigación Operativa y las XII Jornadas de Estadística Pública.
25. Hensher, D.A. and Ton, T.T. (2000). *A comparison of the predictive potential of artificial neural networks and nested logit models for commuter mode choice*. Transportation Research Part E: Logistics and Transportation Review. Volume 36, Issue 3, September 2000, Pages 155-172.
26. Karlaftis, M.G. and Vlahogianni, E.I. (2011); *Statistical methods versus neural networks in transportation research: Differences, similarities and some insights*. Transportation Research Part C: Emerging Technologies. Volume 19, Issue 3, June 2011, Pages 387-399

27. Nijkamp, P., Reggiani, A., Tsang, W.F. ., (2004). *Comparative modelling of interregional transport flows: Applications to multimodal European freight transport*. European Journal of Operational Research. Volume 155, Issue 3, 16 June 2004, Pages 584-602
28. Holland, A. (2003). *Bayesian networks for transport decision scenarios*. Dpt. Computer Science, University of Dortmund.
29. Horn, Mark E.T. (2003). *An extended model and procedural framework for planning multi-modal passenger journeys*. Transportation Research Part B 37, pp. 641-660
30. Jensen, F.V. (2001) *Bayesian Networks and Decision Graphs*. Barcelona. Springer
31. Korb, K., Nicholson, A. (2004). *Bayesian Artificial Intelligence*. Boca Ratón, FL: Chapman and Hall
32. Lafferty, J., Liu, H. and Wasserman, L. (2012). *Sparse nonparametric graphical models*. Statistical Science, 27, 519-537.
33. Lancaster, K. J. (1966). *A New Approach to Consumer Theory*. J. Polit. Econ., vol. 74, no. 2, pp. 132–157, Apr. 1966.
34. Lauritzen, S. L., Dawid, A. P., Larsen, B. N. y Leimer, H. G. (1990). *Independence Properties of Directed Markov Fields*. Networks. 20, 491—505.
35. Lauritzen, S. L. (1992). *Propagation of probabilities, means and variances in mixed graphical association models*. Journal of the American Statistical Association, 87, 1098-1108.
36. Lauritzen, S. L. (1996). *Graphical Models*. Oxford: Clarendon Press.
37. Lauritzen, S. L. y Jensen, F. (2001). *Stable local computation with conditional Gaussian distributions*. Statistics and Computing, 11, 191—203.
38. Lippman, R.P. (1987). *An Introduction to Computing with Neural Nets*. IEEE ASSP Magazine 4, 4-22.

39. Liu, H., Lafferty, J. and Wasserman, L. (2009). *The Nonparanormal: Semiparametric estimation of high dimensional undirected graphs*. J. Mach. Learn. Res. 10, 2295-2328. MR2563983
40. Moral, S., Rumí, R., Salmerón, A. (2001). *Mixtures of truncated exponentials in hybrid Bayesian networks*. Lecture Notes in Artificial Intelligence, Vol. 2143, pp.135-143.
41. Neapolitan, Richard E. (1989). *Probabilistic reasoning in expert systems: theory and algorithms*. Wiley. ISBN 978-0-471-61840-9
42. Neapolitan, R.E. (2004). *Learning Bayesian Networks*. Prentice Hall.
43. Nogal-Macho, M. (2011). *Métodos matemáticos para la predicción de tráfico*. ISBN: 978-84-86116-55-2.
44. Normand, S.L. y Tritchler, D. (1992). *Parameter Updating in Bayes Network*. Journal of the American Statistical Association, 87, 1109—1115.
45. Novikoff, A.B. (1962). On convergence proofs on perceptrons. Symposium on the Mathematical Theory of Automata (pp. 615-622). Polytechnic Institute of Brooklyn.
46. Ortúzar, J., y Willumsen, L. G. (1995). *Modelos de Transporte*. Universidad de Chile.
47. Ortúzar, J.D., y Willumsen, L. G. (2011). *Modelling Transport*. Edited by John Wiley&Sons. 4th ed. Chichester, West Sussex.
48. Pearl, J. (1985). *Bayesian Networks: A Model of Self-Activated Memory for Evidential Reasoning* (UCLA Technical Report CSD-850017). Proceedings of the 7th Conference of the Cognitive Science Society, University of California, Irvine, CA. pp. 329–334
49. Pearl, J. (1986). *A constraint-propagation approach to probabilistic reasoning*. Proceedings of American Association for Artificial Intelligence National Conference on AI, Pittsburgh, Pennsylvania, pp. 133-136.

50. Pearl, J (1988). *Probabilistic Reasoning in Intelligent Systems*. San Francisco CA: Morgan Kaufmann. p. 1988. ISBN 1558604790
51. Riedmiller M. and Braun H. (1993). *A direct adaptive method for faster back-propagation learning: The RPROP algorithm*. Proceedings of the IEEE International Conference on Neural Networks (ICNN), pages 586-591. San Francisco.
52. Russell, S.J., Norvig, P. (2009). *Artificial Intelligence: A Modern Approach*. Boca Raton, FL: Prentice Hall, 3rd edition.
53. Rios-Prado, R., Crespo-Pereira, D., Del Rio-Vilas , D., Rego-Monteil , N., and De Gregorio-Vicente, Ó. (2013). *Model development for the assessment of an international railway corridor – methodological overview*. Proceedings 15th International Conference on Harbor, Maritime and Multimodal Logistics Modelling and Simulation, pp. 133–139. 2013.
54. Salinas R. (2013). Reunión Regional preparatoria de la Conferencia de Exmen Global Decenal de la Ejecución del Programa de Acción de Almaty. Recuperado de: <https://slideplayer.es/slide/2763492/>
55. Sánchez-Cambronero García-Moreno, S. (2008). *Traffic prediction models using bayesian networks and other tools*.
56. Scutari, M. y Denis, J.B. (2015). *Bayesian Networks, with examples in R*. CRC Press, Taylor & Francis group. A Chapman & Hall book, pp. 37-52.
57. Shenton, L.R. and Bowman, K.O. (1977). *A bivariate model for the distribution of $\sqrt{b_1}$ and b_2* . Journal of the American Statistical Association, vol 72, 206-211.
58. Sun, S., Zhang, Ch. and Yu, G., (2006). *A bayesian network approach to traffic flow forecasting*. IEEE Transactions on Intelligent Transportation Systems, vol. 7, no. 1, pp. 124-132. DOI: 10.1109/TITS.2006.869623.
59. Sklar, M. (1959). *Fonctions de répartition a n dimensions et leurs marges*. Publ. Inst. Statist. Univ. Paris 8 229-231. MR0125600

60. Spear, B.D. (1977). *Applications of new travel demand forecasting techniques to transportation planning: a study of individual choice models*. Washington D.C. : Federal Highway Administration, U.S. Department of-Transport, Office of Highway Planning.
61. Tai-Yu, M. (2015). *Bayesian Networks for Multimodal Mode Choice Behavior Modelling: A Case Study for the Cross Border Workers of Luxembourg*. Transportation Research Procedia, vol. 10, pp. 870-880.
62. Tang, D. N., Yang, M., & Zhang, M. H. (2012). *Travel Mode Choice Modeling: A Comparison of Bayesian Networks and Neural Networks*. Applied Mechanics and Materials, 209–211, 717–723. <https://doi.org/10.4028/www.scientific.net/amm.209-211.717>.
63. Tsamardinos, I., Aliferis, C. F. and Statnikov, A. (2003). *Time and sample eficiente discovery of Markov blankets and direct causal relations*. Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 673-678.
64. Tsamardinos, I., Brown, LE., Aliferis, CF. (2006). *The max-min hill-climbing Bayesian network structure learning algorithm*. Machine Learning, 65:31-78.
65. Viviani García, P., Gómez-Villegas, M.A., Maín-Yaque, P. (2014). *Análisis de sensibilidad a la evidencia en Redes Bayesianas Gaussianas*.
66. Weidl, G., Madsen, A. L., Wang, S., Kasper, D. y Karlsen, M., *Early and Accurate Recognition of Highway Traffic Maneuvers Considering Real World Application: A Novel Framework Using Bayesian Networks*. IEEE Intelligent Transportation Systems Magazine, vol. 10, no. 3, pp. 146-158, Fall 2018, doi: 10.1109/MITS.2018.2842049.
67. Williams, H. C. W. L. (1977). *On the formation of travel demand models and economic evaluation measures of user benefit*. Environ. Plan. A, vol. 9, no. 3, pp. 285–344.

-
68. Wilson, E. B. and Hilferty, M. M. (1931). *The distribution of chi-square*. Proceedings of the National Academy of Sciences of the United States of America, 17, 684-688.
 69. Yannis Tyrinopoulos, Constantinos Antoniou (2013). *Factors affecting modal choice in urban mobility*. European Transport Research Review, Volume 5, Issue 1, pp 27–39.
 70. Zhao, T., Liu, H., Roeder, K., Lafferty, J. and Wasserman, L. (2012). *The huge package for high-dimensional undirected graph estimation in R*. The Journal of Machine Learning Research, 98888:1059-1062. R package.
<http://cran.rproject.org/web/packages/huge/index.html>

Apéndice A

Muestra de la Base de Datos

Si bien los datos originales no pueden ser accesibles por razones de confidencialidad asociadas al Caso de Estudio, una muestra de los datos utilizados en el presente trabajo puede descargarse de la siguiente dirección personal: <http://www.mat.ucm.es/~villegas/miembro-cv-oscar.html>. El fichero generado está protegido, por lo que se deberá solicitar el permiso correspondiente para su utilización.

Apéndice B

DATASET ELECCION MODAL

A continuación, se incluye una muestra procedente del archivo de datos “Dataset Elección Modal” para observar su estructura.

Apéndice C

RDATASET: DATASET

A continuación, se incluye una muestra de los datos procedentes del fichero Excel “Dataset Eleccion Modal”, normalizados al rango (-50, 800) en el caso de los tiempos, y al rango (-50, 500) en el caso de los costes. Este fichero contiene los siguientes grupos de columnas en orden de izquierda a derecha:

1. Encabezados gris: tipos de mercancía / ID por OD, precio FOB, TMPA.
2. Encabezados azul: incremento en tiempo respecto a la mejor alternativa.
3. Encabezados naranja: Incremento en coste respecto a la mejor alternativa.
4. Encabezados azul: incremento en tiempo respecto a la mejor alternativa, normalizado al rango (-50, 800).
5. Encabezados naranja. incremento en coste respecto a la mejor alternativa normalizado al rango (-50, 500).
6. Encabezado azul oscuro. Porcentajes de reparto modal (valor fundamental para realizar la comparativa entre los distintos métodos MNL, RN y RB).

Apéndice D

RDATASET: DATANORM

A continuación, se incluye una muestra procedente del archivo de datos para observar su estructura.

Apéndice E

MACRO: R_Red Neuronal.R

```
#Red Neuronal
#Leo la tabla de datos de elección modal
elecmod <- read.delim(C:/R/rrnn/elecmod.txt")
#Modelo de la RRNN
fff <- E1+E2+E3+E4+E5+E6+E7+E8+E9+E10+E11+E12
      ~ FOBPrice+MTPA+T1+T2+T3+T4+T5+T6+T7+T8+T9+T10+T11+T12+T13+T14+
      C1+C2+C3+C4+C5+C6+C7+C8+C9+C10+C11+C12+C13+C14
elecmod$random <- runif(456)
#Muestra de ajuste: 80 % de la muestra
trainset <- subset(elecmod, random <= 0.8)
tt <- trainset
computeSet <- data.frame(FOBPrice = elecmod$FOBPrice, MTPA = elecmod$MTPA, T1 = elecmod$T1, T2 = elecmod$T2, T3 =
elecmod$T3, T4 = elecmod$T4, T5 = elecmod$T5, T6 = elecmod$T6, T7 = elecmod$T7, T8 = elecmod$T8, T9 = elecmod$T9,
T10 = elecmod$T10, T11 = elecmod$T11, T12 = elecmod$T12, T13 = elecmod$T13, T14 = elecmod$T14, C1 = elecmod$C1, C2=
elecmod$C2, C3 = elecmod$C3, C4 = elecmod$C4, C5 = elecmod$C5, C6 = elecmod$C6, C7 = elecmod$C7, C8 = elecmod$C8, C9
= elecmod$C9, C10 = elecmod$C10, C11 = elecmod$C11, C12 = elecmod$C12, C13 = elecmod$C13, C14 = elecmod$C14, E1 =
elecmod$E1, E2 = elecmod$E2, E3 = elecmod$E3, E4 = elecmod$E4, E5 = elecmod$E5, E6 = elecmod$E6, E7 = elecmod$E7, E8 =
elecmod$E8, E9 = elecmod$E9, E10 = elecmod$E10, E11 = elecmod$E11, E12 = elecmod$E12)
computeSet$random <- runif(456)
trainset <- subset(computeSet, random <= 0.8)
computeSet2 <- computeSet[, 1:30]
#Ajuste de los modelos de redes
nn2 <- neuralnet(fff, trainset, hidden= c(2), rep=1, lifiesign = "full", stepmax=1e6, threshold = 0.01)
rr <- compute(nn2, computeSet2)
write.table(netSol$response, C:/R/rrnn/netSol2.txt")
nn4 <- neuralnet(fff, trainset, hidden= c(4), rep=1, lifiesign = "full", stepmax=1e6, threshold = 0.01)
rr <- compute(nn4, computeSet2)
write.table(netSol$response, C:/R/rrnn/netSol4.txt")
nn6 <- neuralnet(fff, trainset, hidden= c(6), rep=1, lifiesign = "full", stepmax=1e6, threshold = 0.01)
rr <- compute(nn6, computeSet2)
write.table(netSol$response, C:/R/rrnn/netSol6.txt")
```

Apéndice F

MACRO: Trainnet.R

```
library(neuralnet)
#Leo la tabla de datos de elección modal
elecmod <- read.delim(C:/R/rrnn/elecmod.txt")
#Modelo de la RRNN
fff <- E1+E2+E3+E4+E5+E6+E7+E8+E9+E10+E11+E12 ~
      FOBPrice+MTPA+T1+T2+T3+T4+T5+T6+T7+T8+T9+T10+T11+T12+T13+T14+
      C1+C2+C3+C4+C5+C6+C7+C8+C9+C10+C11+C12+C13+C14
elecmod$random <- runif(456)
#trainset <- subset(elecmod, random <= 0.8)
#tt <- trainset
computeSet <- data.frame(FOBPrice = elecmod$FOBPrice, MTPA = elecmod$MTPA, T1 = elecmod$T1, T2 = elecmod$T2, T3 =
elecmod$T3, T4 = elecmod$T4, T5 = elecmod$T5, T6 = elecmod$T6, T7 = elecmod$T7, T8 = elecmod$T8, T9 = elecmod$T9,
T10 = elecmod$T10, T11 = elecmod$T11, T12 = elecmod$T12, T13 = elecmod$T13, T14 = elecmod$T14, C1 = elecmod$C1, C2=
elecmod$C2, C3 = elecmod$C3, C4 = elecmod$C4, C5 = elecmod$C5, C6 = elecmod$C6, C7 = elecmod$C7, C8 = elecmod$C8, C9
= elecmod$C9, C10 = elecmod$C10, C11 = elecmod$C11, C12 = elecmod$C12, C13 = elecmod$C13, C14 = elecmod$C14, E1 =
elecmod$E1, E2 = elecmod$E2, E3 = elecmod$E3, E4 = elecmod$E4, E5 = elecmod$E5, E6 = elecmod$E6, E7 = elecmod$E7, E8 =
elecmod$E8, E9 = elecmod$E9, E10 = elecmod$E10, E11 = elecmod$E11, E12 = elecmod$E12)
#Muestra de ajuste: 80 % de la muestra
computeSet$random <- runif(456)
trainset <- subset(computeSet, random <= 0.8)
crossValSet <- subset(computeSet, random > 0.8)
computeSet2 <- computeSet[, 1:30]
crossValSet2 <- crossValSet[, 1:30]
#Ajuste de los modelos de redes
nn2 <- neuralnet(fff, trainset, hidden= c(2), rep=1, lifesign = "full", stepmax=1e6, threshold = 0.01)
rr <- compute(nn2, crossValSet2)
write.table(nn2$response, C:/R/rrnn/netSol2.txt")
nn4 <- neuralnet(fff, trainset, hidden= c(4), rep=1, lifesign = "full", stepmax=1e6, threshold = 0.01)
rr <- compute(nn4, computeSet2)
write.table(netSol$response, C:/R/rrnn/netSol4.txt")
nn6 <- neuralnet(fff, trainset, hidden= c(6), rep=1, lifesign = "full", stepmax=1e6, threshold = 0.01)
rr <- compute(nn6, computeSet2)
write.table(netSol$response, C:/R/rrnn/netSol6.txt")
```

Apéndice G

MACRO: R_Red Bayesiana.R

```
#Redes Bayesianas
library(bnlearn)
#Test previos y normalización de la muestra
dh <- DH.test(computeSet,Y.names=names(computeSet))
tt <- DH.test(computeSet)
C <- huge.npn(computeSet)
C <- as.data.frame(C)
tt <- DH.test(C)
tt
#Ajuste de la red bayesiana
rbg <- mmhc(C, blacklist = blackl)
bns <- arc.strength(rbg, C)
bns <- boot.strength(C)
bnf <- bn.fit(rbg, C)
bnstruct <- bn.net(bnf)
#Generacion de la muestra con las predicciones
pp1 <- predict(bnf, .E1", C)
pp <- as.data.frame(pp1)
pp$pp1 <- pp$pp1 - C$E1
pp2 <- predict(bnf, .E2", computeSet)
pp$E2 <- pp2 - C$E2
pp3 <- predict(bnf, .E3", computeSet)
pp$E3 <- pp3 - C$E3
pp4 <- predict(bnf, .E4", computeSet)
pp$E4 <- pp4 - C$E4
pp$E5 <- predict(bnf, .E5", computeSet) - C$E5
pp$E6 <- predict(bnf, .E6", computeSet) - C$E6
pp$E7 <- predict(bnf, .E7", computeSet) - C$E7
pp$E8 <- predict(bnf, .E8", computeSet) - C$E8
pp$E9 <- predict(bnf, .E9", computeSet) - C$E9
pp$E10 <- predict(bnf, .E10", computeSet) - C$E10
pp$E11 <- predict(bnf, .E11", computeSet) - C$E11
pp$E12 <- predict(bnf, .E12", computeSet) - C$E12
write.table(pp, C:"\R\rrnn\rby.txt")
```