



TRABAJO FIN DE GRADO

Un modelo para el diseño robusto de redes y análisis
estadístico de su eficiencia.

Realizado por: Lorena Rodríguez González

Supervisado por: Eduardo Conde Sánchez

FACULTAD DE MATEMÁTICAS
DPTO DE ESTADÍSTICA E INVESTIGACIÓN OPERATIVA

Índice

1. Introducción	6
1.1. Modelando problemas logísticos mediante grafos.	6
1.2. Algunas aplicaciones reales.	8
1.2.1. Mejora de la eficiencia de la red eléctrica.	8
1.2.2. Interconexión de puertos para hacerlos más operativos.	9
1.2.3. Infraestructura robusta en el sector de la industria.	11
1.2.4. Conectividad robusta frente a catástrofes.	12
1.2.5. Producción y transporte de gas en España.	13
2. Formulación matemática del problema.	15
2.1. Diseño robusto bajo costes inciertos de transporte y mantenimiento.	15
2.1.1. Elementos del problema.	15
2.1.2. Formulación del problema determinístico.	16
2.1.3. Ejemplo del problema (DUND)	18
2.1.4. Formulación del problema minmax regret.	19
2.1.5. Ejemplo (DUND) (<i>continuación</i>)	25
2.2. Diseño robusto bajo costes inciertos y pares vinculados.	27
2.2.1. Formulación del problema de separación y maestro.	27
2.2.2. Ejemplo (DUND) (<i>continuación</i>)	32
2.3. Implementación algorítmica mediante AMPL	33
2.3.1. Código AMPL del problema maestro (MP) y del de separación (SP)	33
2.3.2. Redes obtenidas durante la aplicación del algoritmo	41
2.3.3. Detalle de la salida del algoritmo	42
3. Estudio experimental.	55
3.1. Experimento con costes independientes	58
3.1.1. Distribución uniforme	58
3.1.2. Distribución Beta	59
3.2. Experimento con costes relacionados.	61

3.2.1. Distribución uniforme discreta.	61
3.2.2. Dependencia de los arcos con distribución uniforme discreta según su orientación.	63
3.2.3. Dependencia de los arcos según una distribución uniforme.	65
3.2.4. Dependencia de los arcos con distribución uniforme según su orientación.	66
3.2.5. Costes dependientes con distribución beta.	67
3.2.6. Dependencia del coste según la orientación del arco.	69
3.3. Conclusiones.	71

Abstract

The aim of this work is to design a robust distribution network under uncertainty in costs. These parameters can be modeled as points in a polyhedral set where the knowledge about dependence relations among them are represented. We will use the minmax regret paradigm that looks for a solution with an efficiency as close as possible to the optimum under each scenario of costs. These uncertain parameters will model the scenario under which the new network will be evaluated.

First, we give an introduction motivating the purpose and usefulness of this problem in a globalized economy. Studies in different areas of the economy will show the need of mathematical models where the existing uncertainty in the different elements of the network is incorporated.

Chapter 2 deals with the deterministic problem that solves a flow problem determining how to transport certain commodities from their origins to their destinations. We will analyse the design of a network with potential arcs under uncertain costs according to find a solution as close as possible to the best one under each scenario. Moreover, we will develop a Benders Decomposition algorithm to find numerically a robust design that solves this problem.

In the last part of this work, different statistical studies have been carried out by using the software RStudio to compare the minmax regret network design with respect to the optimal design under the expected costs for a family of distributional conditions.

1. Introducción

Los primeros avances relevantes en el análisis de problemas logísticos mediante algoritmos de optimización podríamos situarnos en la década de los años 50. El algoritmo de *Ford Fulkerson* significó un método de determinación del máximo flujo de producto que admitía una red de comunicaciones. Desde entonces han aparecido multitud de estudios debido a la importancia del problema tratado (véase: RK Ahuja, TL Magnanti, JB Orlin (1988), *Network flows*.)

En este trabajo se modela el problema de construcción o mantenimiento de una red logística junto con el de determinación de un flujo óptimo de mercancías bajo un escenario de costes inciertos. Usaremos como paradigma el criterio *minmax regret* y denominaremos robusto al diseño óptimo resultante bajo dicho criterio.

1.1. Modelando problemas logísticos mediante grafos.

La incertidumbre en los parámetros continuos se modela a través de conjuntos poliédricos que combinaremos con otro tipo de incertidumbre de naturaleza discreta. El grado de conocimiento sobre estos parámetros se debe en algunos casos, a la falta de datos para realizar estimaciones mientras que en otros casos si bien existen datos, podrían volverse obsoletos debido a la modificación que produciría la creación de la nueva infraestructura que se está diseñando.

En la práctica es frecuente encontrar múltiples diseños óptimos bajo las mismas condiciones y en esta situación se podría elegir uno con características adecuadas desde el punto de vista de la robustez. Por tanto una estrategia razonable podría ser buscar una solución cuasi-óptima en una variedad de escenarios posibles. Utilizaremos el paradigma *minmax regret* con respecto a un conjunto dado de parámetros inciertos que modelan el escenario bajo el cuál se evaluará la nueva red. Se tratará de buscar una solución con una eficiencia lo más cercana posible a la óptima bajo cada escenario. Además el problema de los costes inciertos se va a modelar utilizando un poliedro general que se integra en la optimización del *minmax regret* y el poliedro de dichos costes se combinará con un conjunto discreto de posibles escenarios, que conllevará ventajas como la facilidad

para identificar las relaciones existentes entre los parámetros desconocidos. Básicamente los parámetros continuos modelan costes inciertos de construcción, mantenimiento o transporte. Los parámetros discretos se refieren a proveedores o clientes del producto suministrado.

Este problema también se ha estudiado en el contexto de la agilidad de fabricación en entornos competitivos en los que surgen oportunidades de mercado y desaparecen sin patrones conocidos cómo podría ser la necesidad de compra a proveedores más baratos pero menos confiables o no probados. La pérdida de proveedores se considerará en nuestro modelo definiendo diferentes escenarios para las ubicaciones disponibles desde dónde la demanda debe ser satisfecha.

Según la naturaleza de la incertidumbre, los estudios existentes se podrían categorizar en tres grupos, estocásticos, robustos y difusos. El diseño de redes de suministro robusto bajo un conjunto continuo de parámetros inciertos se encuentra en el segundo grupo y el hecho de usar poliedros en lugar de hipercubos para modelar la incertidumbre proporciona al modelo una mejor aproximación a situaciones reales.

Desde el punto de vista computacional el diseño de una red de suministro real tiene una gran complejidad debido al gran tamaño de los problemas y a la no linealidad de las restricciones. Además en presencia de parámetros inciertos, el problema de optimización resultante es todavía más complejo. De hecho, incluso bajo elecciones muy simples del conjunto de posibles demandas, la complejidad del problema se vuelve NP -dura.

El rediseño de redes es un problema actual, importante en una economía globalizada. Vamos a considerar dicho problema en varios contextos. Por ejemplo el ámbito de la conexión entre puertos, dónde puede existir una desorganización en el transporte de mercancías que provoca el aumento de los costes. Otro ejemplo es en el ámbito de la industria, ya que las condiciones extremas ralentizan las maquinarias y el criterio *minmax regret* nos ayudará a optimizar su funcionamiento. Algunos tipos de incertidumbre que modelaremos en el siguiente capítulo podrían ser debidos a situaciones excepcionales cómo una pandemia, una guerra o una catástrofe natural, por ello en el ámbito empresarial, es interesante buscar soluciones robustas que estén preparadas para cualquier

escenario posible dentro de lo razonable.

1.2. Algunas aplicaciones reales.

1.2.1. Mejora de la eficiencia de la red eléctrica.

Un estudio del *Instituto de la energía del University College* de Dublín publicado en el periódico *El Confidencial* [2] sugiere que podría ahorrarse hasta un 48% de los costes energéticos modificando el actual modelo de red eléctrica. En el estudio se propuso una nueva red que conecte todos los países europeos sin restricciones ya que las interconexiones incrementan la eficiencia de los sistemas de transmisión porque permiten un intercambio de energía desde dónde es más barata hasta dónde es más cara en cada momento.

Se analizaron tres modelos de red eléctrica: el actual, el paneuropeo y otro que conecte todos los países europeos sin restricciones, con un objetivo claro: acabar con los sistemas aislados y promover un suministro eléctrico más seguro. Si al modelo actual le incorporásemos más energías renovables podría derivar en una sobrecarga de la red, por lo que los costes aumentarían. Aunque existen motivos para pensar que el nuevo modelo de interconexión entre países ahorraría hasta un 48% de los costes, no se cuenta con datos que permitan respaldarlo por lo que se debe tratar la situación asumiendo incertidumbre en los costes.

Según los datos del periódico *Cotizalia* [3], en España la interconexión eléctrica es muy escasa y de hecho está por debajo del mínimo establecido en la Unión Europea. Esto es debido a que España depende de Francia para estar conectada y todas las redes deben atravesar dicho país antes de llegar a España. El problema viene porque Francia apenas alcanza un 3% de interconexión con el resto de países lo que hace que España tenga un primitivo modelo de red eléctrica.

Este posible cambio de la red supondría una gran inversión a la que habría que añadir la incertidumbre existente sobre los beneficios que permitirían recuperarla. Por esta razón sería necesario una optimización de las ubicaciones de los centros de producción de energías renovables disminuyendo así la capacidad total instalada. *Eddie O' Connor*

suministro. También existe un desequilibrio entre oferta y demanda, los barcos se saltan escalas, los puertos suspenden su actividad y se producen otros fenómenos que añaden incertidumbre al problema.

Los puertos de Estados Unidos y China, por los que pasa la mayor cantidad de mercancía, se encuentran colapsados y esto se debe, además de la escasez de recursos, a la primitiva conexión entre sus instalaciones.

Las piezas de reparación, los barcos y los contenedores no se encuentran dónde se necesitarían o esperábamos encontrar y ahí tenemos una nueva fuente de incertidumbre. Se produce lo que se conoce como efecto látigo, cuando hay escasez de recursos la gente compra mucho más (caso de demanda incierta). Las empresas hacen pedidos que se retrasan o no llegan, cómo hay menos barcos las empresas se quedan sin stock, en consecuencia piden más y tratan de almacenar masivamente. Por tanto además de la incertidumbre existen relaciones de dependencia entre los diferentes elementos que determinan los costes (esto lo comentaremos de nuevo en el capítulo 3 dedicado a realizar un estudio experimental del modelo matemático propuesto).

Otro aspecto a destacar es el solapamiento de diferentes redes logísticas. Por ejemplo, Estados Unidos ha llegado a tener un gran atasco en sus puertos por la entrada de más de 7.000 buques para descargar contenedores. Como medida principal para solucionarlo, se estableció un horario completo de trabajo, lo que derivó dicho atasco a otro eslabón de la cadena ya que los puertos son sólo la primera parada en el camino del transporte de la mercancía. La situación se agrava debido a que la disponibilidad de camiones para mover mercancías no coinciden con las horas ampliadas en los puertos.

Algunos de estos retrasos y desequilibrios son debidos a los efectos de la pandemia *COVID – 19* pues supuso un parón de demanda, la disfuncionalidad de muchos barcos hizo que la logística marítima no estuviese preparada para una recuperación tan rápida y desigual. Según un estudio mencionado en el periódico *elDiario.es* [4] sería posible proponer un nuevo modelo que cambie la línea logística tradicional, flexibilizándose el proceso de fletar barcos con el objetivo de disminuir los costes y mejorar las conexiones entre puertos haciéndolos más operativos. Así se disminuiría los largos viajes de los bar-

cos y por tanto dejarían de saltarse escalas. En cualquier caso, estos estudios ponen de manifiesto la necesidad de modelar matemáticamente el problema mediante un paradigma que permita incorporar la incertidumbre existente en los diferentes elementos de la red.

1.2.3. Infraestructura robusta en el sector de la industria.

Entre las partes más importantes del desarrollo de la industria digital se encuentra la incorporación de estructuras robustas de cableado que permiten a las empresas gestionar de forma inteligente la maquinaria a través del control remoto. La infraestructura de redes lo es todo ya que, a través de ella, se conectan y comunican todos los dispositivos que permiten la automatización. Según un estudio publicado en *Cluster Industrial* [6] el 80 % de las empresas manufactureras hacen uso de la conectividad inalámbrica para el control y automatización de la planta industrial, un sistema de cableado robusto que permite una conexión inalámbrica estable.

Existe hoy en día una gran presión para implementar de forma adecuada nuevas estrategias de inteligencia artificial mediante conectividad inalámbrica para la automatización de maquinaria. Sin embargo, a la hora de implementar estas iniciativas se enfrentan a un desafío fundamental: la tecnología y la infraestructura necesarias para sostenerlas son excepcionales y muy complejas. El consumo y transporte de datos requiere una nube segura y muy eficiente. No existen datos históricos que puedan ser usados para la toma de decisiones puesto que es una nueva tecnología.

Otro factor que debe considerarse es el de la fiabilidad de la red. Pueden darse condiciones extremas cómo una alta vibración, líquidos corrosivos o intolerancia, en las que a pesar de ello será necesario una conexión estable. La dificultad de conseguir una potente conexión inalámbrica durante las condiciones adversas provoca un incremento de los costes. La digitalización es fundamental en este sector pues optimiza la comunicación industrial y mantiene el tiempo de actividad de la red administrando seguridad a la maquinaria. Una estructura robusta de la red reduciría los costes de producción y la redundancia además de optimizar el rendimiento conjunto de la planta. No obstante, el

diseño de la red debe incorporar todos los elementos que se han comentado anteriormente y conseguir que la eficiencia, incluso en las peores condiciones, sea lo mejor posible. Este es el propósito del criterio *minmax regret* que desarrollaremos en el siguiente capítulo.

1.2.4. Conectividad robusta frente a catástrofes.

La reciente pandemia y hechos cómo la guerra de Ucrania están poniendo al descubierto las carencias de las empresas en cuánto al flujo de información y al trabajo en remoto. La actividad económica actual hace que este factor de conectividad condicione cada vez más el éxito o fracaso de la empresa. También las compañías españolas demandan acceso a la nube, así como conexiones robustas, seguras y fiables entre clientes privados, sucursales y empleados, es decir, conectar los centros de datos e incrementar la seguridad entre accesos. Según el estudio de *Redes&Telecom* [5] las empresas españolas están teniendo en cuenta estos factores para imponer nuevos requisitos de conectividad. *Colt* es una compañía que provee una red mundial única integrada con el sistema de la nube que tiene una oferta y demanda que permite más automatización y menos intervención local. La multinacional quiere establecer una infraestructura de fibra que tenga la capacidad de llegar a todas las urbes más importantes en todas las ciudades.

Esta iniciativa de infraestructura ha sido principalmente motivada por la reciente pandemia pues la demanda de los consumidores era incierta, por ejemplo se esperaba que 2020 fuese un año de crecimiento y finalmente ha sido un año de adaptaciones. La demanda de los consumidores de soluciones digitales fue mucho mayor que en años anteriores, lo que hizo que las empresas no pudieran satisfacer toda la demanda. En dichos meses, *Colt* aumentó entre un 30 y un 40% su capacidad normal de trabajo a través del teletrabajo, pues la compañía no era ajena a esta modalidad de funcionamiento en remoto. La empresa también abrió nuevas localizaciones físicas, amplió las comunicaciones mediante cable submarino y adecuó el uso de la nube. En definitiva, constituye un ejemplo práctico de cómo las empresas deber ser ágiles en la adaptación de sus infraestructuras para garantizar la conectividad ante escenarios impredecibles como pandemias, guerras y otros tipos de catástrofes.

1.2.5. Producción y transporte de gas en España.

La técnica "fracking" consiste en la extracción de recursos fracturando rocas que contienen petróleo y gas en su interior. En un reciente artículo del periódico *La Razón* [7] se destaca que la eficiencia de la producción de gas dependerá de dicha técnica y de la demanda en las ciudades, aunque en ambos casos existen importantes factores inciertos. En 2014, mientras que esta técnica estaba de moda en Europa, en España había un centenar de permisos exploratorios para sondear gas y petróleo en las principales cuencas mineras. En este sondeo se averiguó que a pesar de que España podría almacenar un tercio del gas que necesita Europa, sigue dependiendo al 99% de las importaciones de hidrocarburos. En el caso de Europa central se agrava la situación por su dependencia del gas ruso. De hecho, en 2014, el gobierno de España estableció un plan de incentivos tributarios y no tributarios para que las comunidades autónomas, ayuntamientos y propietarios tuvieran un beneficio del nuevo impuesto sobre el valor de la producción del petróleo y gas. Además si la explotación de hidrocarburos se hace en una zona cercana al mar se beneficiaría la comunidad autónoma más próxima al yacimiento y las empresas con autorización de explotación establecerían un retorno del 1% del valor de la producción anual para los propietarios de los terrenos explotados.

España es un país semiexplorado en materia de hidrocarburos. Además los riesgos sísmicos que se tendrían como consecuencia de la técnica *fracking* son irrelevantes ya que la energía liberada durante la técnica es muy inferior a la liberada en una mina. Existen experiencias pioneras como las de Estados Unidos que se ha convertido en el mayor exportador del gas natural gracias a su fracturación hidráulica y su incremento de la producción. Debido a esto, Europa empezó a sondear los recursos fósiles que nunca habían sido explorados en busca de gas y petróleo. Sin embargo, la caída de los precios del gas fue un lastre para la extensión de la técnica *fracking* en Europa y derivó en su prohibición en muchas comunidades autónomas. Todo esto ha cambiado en el nuevo escenario geopolítico en el que se buscan alternativas al gas ruso como la importación de GNL estadounidense, un 40% más caro que el ruso debido a la logística necesaria para el traslado y tratamiento.

Existen estudios recientes que revelan que España tiene una reserva de gas que en un escenario intermedio supondría unas 70 veces su demanda actual. Se prevé que en 2040 el impacto sobre el PIB de los recursos propios de hidrocarburos equivaldría a una cantidad mucho superior que el de otras actividades como podría ser el transporte o el sector primario. España alberga tanto gas que con una probabilidad de extracción del 50 % tendría un beneficio de hasta 4 veces mayor que con el valor del petróleo. Durante más de 20 años España dispondría de un saldo neto de exportador de gas y reduciría las importaciones nacionales, con lo que se generaría una gran cantidad de puestos de trabajos directos e indirectos.

Aunque en el artículo sólo se dan datos estimados de las ganancias que se podrían generar en España por la explotación de gas, en realidad existe cierta incertidumbre principalmente por la demanda real que se tendría en las ciudades debido a la competencia. Por otra parte, la técnica *fracking* podría ser que no encontrase suficiente gas como es de esperar y que no llegara a satisfacer la demanda. Ambos casos son ejemplos de incertidumbre que habrá que tendremos en cuenta en el próximo capítulo a la hora de modificar y/o ampliar la actual red de gasoductos de acuerdo a los escenarios de posibles yacimientos según se ilustra en la figura 2.

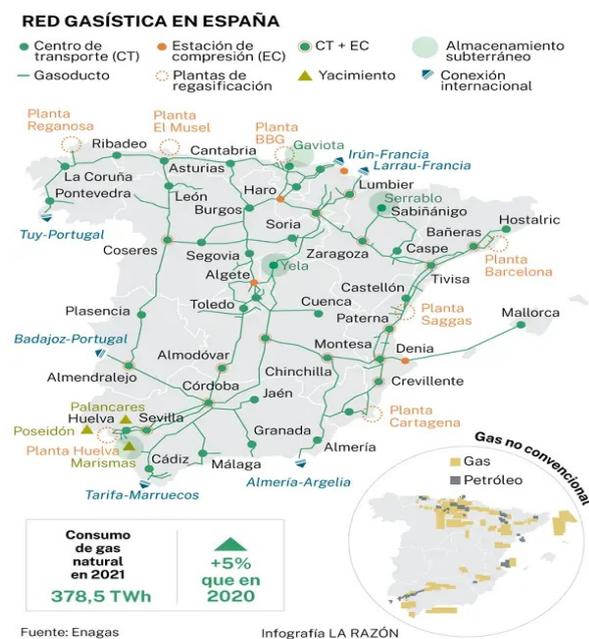


Figura 2: Red gasística en España.

2. Formulación matemática del problema.

En el capítulo anterior se recogían diferentes situaciones reales que aconsejan poder disponer de técnicas para la planificación y diseño de redes logísticas que habrán de funcionar bajo escenarios inciertos. En el presente capítulo plantearemos el problema de optimización bajo diferentes hipótesis sobre el grado de conocimiento existente sobre los diferentes elementos del problema. En primer lugar, estudiaremos el problema de cómo transportar un producto determinado desde origen a destino bajo diferentes escenarios económicos. Presentaremos un modelo con arcos potenciales, es decir, tendremos que decidir si construir dichos arcos para transportar las unidades dependiendo de los costes de construcción y de mantenimiento que en este caso son inciertos.

Inicialmente ilustramos el problema con un modelo determinístico que se irá retomando a lo largo del capítulo para añadir condiciones más complejas que modelan distintos grados de incertidumbre. La parte final del capítulo se dedica al estudio de un esquema de Descomposición de Benders para resolver uno de los modelos propuestos.

2.1. Diseño robusto bajo costes inciertos de transporte y mantenimiento.

A continuación definimos los elementos básicos de un modelo de diseño robusto. Lógicamente representa una situación simplificada que tiene como objetivo mostrar las propiedades del problema de optimización y desarrollar técnicas numéricas de obtención de la solución óptima.

2.1.1. Elementos del problema.

Primero vamos a describir el modelo considerando:

1. $N = \{1 \dots n\}$ un conjunto de nodos.
2. $A \subseteq N \times N$ un conjunto de arcos potenciales que pueden ser, o no, construidos.
3. $P \subseteq N \times N$ una lista de pares de nodos origen-destino que deben estar conectados en la red. Puede ser interpretado cómo enviar una unidad de producto entre ambos

nodos (origen y destino).

4. $f_{ij}^s, \forall (i, j) \in A$ un coste fijo no negativo correspondiente a la construcción o al mantenimiento del arco potencial (i, j) bajo un escenario dado $s \in S$.
5. $c_{ij}^s, \forall (i, j) \in A$ un coste de transporte unitario correspondiente al flujo de las mercancías a través del arco (i, j) bajo el escenario $s \in S$. Estos costes modelan el precio de enviar una unidad del producto, considerado a través de la red (electricidad, combustible).

Nuestro objetivo es encontrar un subconjunto de arcos potenciales $A_y \subseteq A$ que deba construirse para conectar de manera óptima el conjunto de pares origen-destino de la red resultante (N, A_y) . Vamos a considerar las siguientes variables de decisión:

1. $y_{ij} \in \{0, 1\} \forall (i, j) \in A$. Es una variable binaria que toma el valor uno si decido construir el arco potencial $(i, j) \in A_y$ de la red y toma el valor 0 en caso contrario.
2. $x_{ijp} \in \{0, 1\} \forall (i, j) \in A$. Es una variable binaria de flujo que toma el valor 1 si elijo el arco $(i, j) \in A_y$ para conectar el par $p \in P$ de la lista de nodos y toma el valor 0 en caso contrario.

2.1.2. Formulación del problema determinístico.

Recordemos que un problema de flujo determina cómo transportar un cierto producto desde el origen hasta su destino sin tener en cuenta los costes fijos o de mantenimiento. Lo interesante de este modelo es cómo conectar los nodos ya que el diseño está fijo. Una formulación de este problema podría ser:

$$D(s) := \min \sum_{(i,j) \in A} c_{ij}^s x_{ijp}$$

$$\text{suje}to \ a : \sum_{(i,j) \in A} x_{ijp} - \sum_{(k,i) \in A} x_{kip} = b_{ip}, i \in N, p \in P,$$

$$x_{ijp} \in \{0, 1\} \quad (i, j) \in A, p \in P$$

En el anterior modelo conocíamos los arcos disponibles para conectar los nodos pero es posible que algunas de esas conexiones (cómo en el ejemplo del gasoducto 1.2.5) deben ser construidos previamente de acuerdo a ciertos costes fijos o de mantenimiento. La formulación de programación lineal entera mixta que veremos a continuación es la siguiente:

$$D(s) := \min \sum_{(i,j) \in A} f_{ij}^s y_{ij} + \sum_{(i,j) \in A} c_{ij}^s x_{ijp}$$

sujeto a:

$$\sum_{(i,j) \in A} x_{ijp} - \sum_{(k,i) \in A} x_{kip} = b_{ip}, i \in N, p \in P, \quad (DUND)$$

$$x_{ijp} \leq y_{ij} \quad (i, j) \in A, p \in P$$

$$y_{ij} \in \{0, 1\} \quad (i, j) \in A$$

$$x_{ijp} \in \{0, 1\} \quad (i, j) \in A, p \in P$$

dónde $b_{ip} = 1, b_{jp} = -1$ si $p = (i, j) \in N \times N$ y $b_{ip} = 0$ en caso contrario. El primer bloque de restricciones son de balance de flujo. Si $b_{ip} = 0$ entonces dicho nodo actúa cómo un punto de trasbordo, en caso contrario una unidad de mercancía se deja en dicho nodo o sale de dicho nodo. Las restricciones binarias de las variables y_{ij} y x_{ijp} pueden ser relajadas pues se trata de un problema con la propiedad unimodular, pero esta propiedad no se conservará para los modelos que construiremos más adelante.

El conjunto factible podría ajustarse mediante la agregación de una serie de restricciones, por ejemplo, para modelar recursos limitados en la construcción de la red. Además podemos considerar los coeficientes b_{ij} y b_{jp} de manera más general, tomando $b_{ij} = b_p, b_{jp} = -b_p \forall p = (i, j) \in P$ y en caso contrario $b_{kp} = 0$, dónde b_p es un número entero positivo. En este caso, b_p modela la cantidad de mercancía transportada de i a j .

En las formulaciones posteriores usaremos la siguiente notación para el coste determinístico:

$$D(x, y, s) := \sum_{(i,j) \in A} f_{ij}^s y_{ij} + \sum_{(i,j) \in A, p \in P} c_{ij}^s x_{ijp}$$

2.1.3. Ejemplo del problema (DUND)

Ahora presentaremos un ejemplo que se irá adaptando para ilustrar los diferentes grados de incertidumbre considerados en el modelo de diseño robusto de una red.

En el ejemplo se ilustra cuatro orígenes (círculos) con sus respectivos destinos (cuadrados) dónde están todos los arcos potenciales posibles y dónde no hay incertidumbre ni tenemos ningún límite en las capacidades del arco. El problema consiste en el diseño de una red que conecte todos los nodos origen-destino a un coste mínimo con la formulación (*DUND*). Los orígenes pueden ser proveedores y los destinos puntos de demandas. Observemos que con frecuencia la solución es un árbol (spanning tree). En el primer caso todos los costes son determinísticos pues se toma como supuesto que tenemos certeza de los costes mientras que en el segundo existe una zona de incertidumbre dónde los costes no están fijos, es decir, no sabemos con absoluta certeza los costes reales sino que podemos llegar a conocer un intervalo en el que se encuentran dichos costes. Dicha zona se representa en el ejemplo como un rectángulo. Cómo no se conocen los costes, la solución del problema (*DUND*) intentará evitar dicho rectángulo para evitar el riesgo que suponen los escenarios desfavorables.

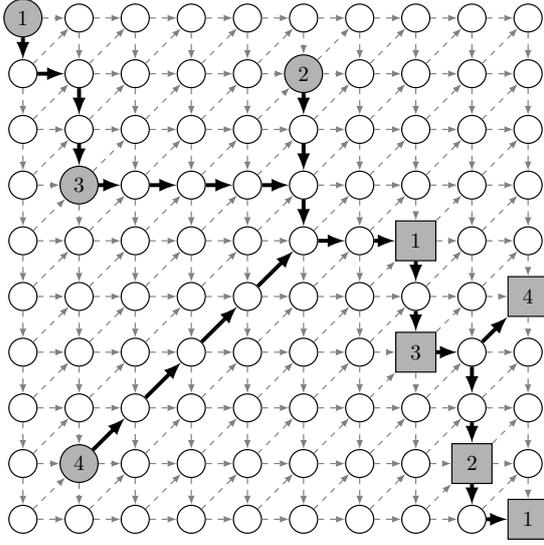


Figura 3: Ejemplo de (DUND).

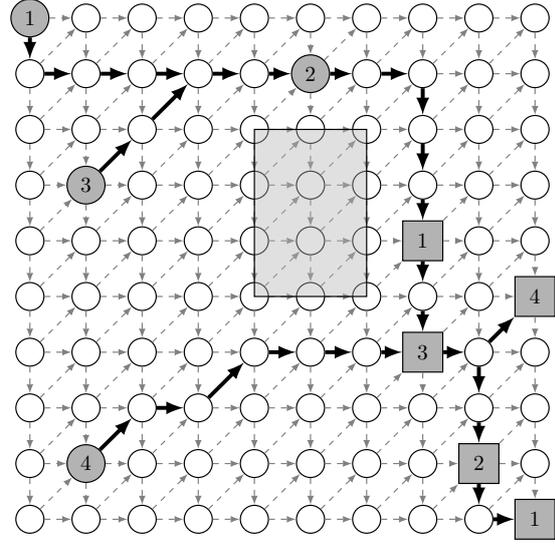


Figura 4: Un diseño robusto evitando la zona de incertidumbre.

2.1.4. Formulación del problema minmax regret.

Sea $R(x, y, s)$ la diferencia positiva bajo un escenario $s \in S$ entre el coste correspondiente al diseño de red y con flujo x y el coste óptimo bajo el escenario s , es decir:

$$R(x, y, s) := \sum_{(i,j) \in A} f_{ij}^s y_{ij} + \sum_{(i,j) \in A} c_{ij}^s x_{ijp} - D(s)$$

dónde $D(s)$ es el valor óptimo de la formulación 2.1.2.

Suponemos que los posibles costes son los pertenecientes a un poliedro S dado por:

$$S := \{(c, f) \geq 0 : \sum_{(i,j) \in A} (a_{ijk}^f f_{ij} + a_{ijk}^c c_{ij}) \leq d_k \quad k \in K\}$$

Los parámetros a_{ijk}^f y a_{ijk}^c son conocidos. Sabemos que un poliedro es convexo y cerrado pero además asumiremos que es compacto por lo que existirán constantes \bar{c} y \bar{f} tales que $0 \leq c_{ij} \leq \bar{c}$, $0 \leq f_{ij} \leq \bar{f}$ con $(c, f) \in S$.

Dado dicho poliedro podríamos calcular $R(x, y)$ que es la máxima diferencia positiva sobre el conjunto de posibles escenarios $s \in S$. Este valor puede obtenerse resolviendo la

siguiente formulación dónde utilizaremos el poliedro anterior cuyas componentes son c_{ij}^s y $f_{ij}^s \forall (i, j) \in A$.

$$R(x, y) := \text{máx} \quad \sum_{(i,j) \in A} f_{ij} y_{ij} + \sum_{(i,j) \in A, p \in P} c_{ij} x_{ijp} \\ - \sum_{(i,j) \in A} f_{ij} \bar{y}_{ij} - \sum_{(i,j) \in A, p \in P} c_{ij} \bar{x}_{ijp}$$

subjeto a:

$$\begin{aligned} \sum_{(i,j) \in A} \bar{x}_{ijp} - \sum_{(k,i) \in A} \bar{x}_{kip} &= b_{ip}, \quad i \in N, p \in P, \\ \bar{x}_{ijp} &\leq \bar{y}_{ij}, \quad (i, j) \in A, p \in P, \\ \sum_{(i,j) \in A} (a_{ijk}^f f_{ij} + a_{ijk}^c c_{ij}) &\leq d_k, \quad k \in K, \\ \bar{y}_{ij} &\in \{0, 1\}, \quad (i, j) \in A, \\ \bar{x}_{ijp} &\in \{0, 1\}, \quad (i, j) \in A, p \in P, \\ f_{ij} &\geq 0, \quad (i, j) \in A, \\ c_{ij} &\geq 0, \quad (i, j) \in A. \end{aligned} \tag{1}$$

Como f_{ij} , \bar{y}_{ij} , c_{ij} , \bar{x}_{ijp} son variables entonces aparecen términos cuadráticos en la función objetivo, lo que resulta un inconveniente para resolver el problema. Para solventarlo dichos términos se pueden linealizar bajo el supuesto de compacidad de nuestro poliedro S . La formulación *MILP* resultante se llama problema de separación (*SP*). Este nuevo problema nos permitirá evaluar un finito y pequeño número de diseños de red para después poder quedarnos con el mejor de todos, pero esto solo será posible cuando tenemos un número suficientemente pequeño. En la práctica el número de opciones es demasiado grande pues estamos interesados en un diseño de red robusto lo que nos lleva a formular el siguiente problema (RND) que selecciona de entre todas las posibles elecciones de subconjuntos de arcos $A_y \subseteq A$ el óptimo.

Para esta nueva formulación hacemos el siguiente cambio de variables con el fin de linealizar la función objetivo teniendo en cuenta que ahora tendremos que hacer cambios

en las restricciones del problema (1).

$$f_{ij}\bar{y}_{ij} = w_{ij}$$

$$c_{ij}\bar{x}_{ijp} = z_{ijp}$$

$$R(x, y) := \text{máx} \quad \sum_{(i,j) \in A} f_{ij}y_{ij} + \sum_{(i,j) \in A, p \in P} c_{ij}x_{ijp} - \\ - \sum_{(i,j) \in A} w_{ij} - \sum_{(i,j) \in A, p \in P} z_{ijp}$$

sujeto a:

$$\begin{aligned} \sum_{(i,j) \in A} \bar{x}_{ijp} - \sum_{(k,i) \in A} \bar{x}_{kip} &= b_{ip}, \quad i \in N, p \in P, \\ \bar{x}_{ijp} &\leq \bar{y}_{ij}, \quad (i, j) \in A, p \in P, \\ \sum_{(i,j) \in A} (a_{ijk}^f f_{ij} + a_{ijk}^c c_{ij}) &\leq d_k, \quad k \in K, \\ c_{ij} - \bar{c}(1 - \bar{x}_{ijp}) &\leq z_{ijp}, \quad (i, j) \in A, p \in P, \\ z_{ijp} &\geq 0, \quad (i, j) \in A, p \in P, \\ f_{ij} - \bar{f}(1 - \bar{y}_{ij}) &\leq w_{ij}, \quad (i, j) \in A, \\ w_{ij} &\geq 0, \quad (i, j) \in A, \\ \bar{y}_{ij} &\in \{0, 1\}, \quad (i, j) \in A, \\ \bar{x}_{ijp} &\in \{0, 1\}, \quad (i, j) \in A, p \in P, \\ f_{ij} &\geq 0, \quad (i, j) \in A, \\ c_{ij} &\geq 0, \quad (i, j) \in A. \end{aligned} \tag{SP}$$

Escribimos el poliedro en forma de restricción para evitar tener que añadir en la función objetivo las respectivas "s" de las variables. Ahora como tenemos el cambio de variable $c_{ij}\bar{x}_{ijp} = z_{ijp}$, si $\bar{x}_{ijp} = 1$ entonces $z_{ij} = c_{ij}$ y efectivamente si $\bar{x}_{ijp} = 0$ entonces $z_{ij} = 0$. En la restricción $c_{ij} - \bar{c}(1 - \bar{x}_{ijp}) \leq z_{ijp}$ con \bar{c} suficientemente grande, como estamos maximizando un sumatorio negativo y $z_{ijp} \geq 0$ si $\bar{x}_{ijp} = 0$ entonces nos va a interesar hacer $z_{ijp} = 0$ que es lo que buscábamos. Además cuando $\bar{x}_{ijp} = 1$ ocurre que

$c_{ij} \leq z_{ijp}$ como maximizamos un sumatorio negativo nos interesa la igualdad $c_{ij} = z_{ijp}$. Al igual podríamos razonar con la restricción $f_{ij} - \bar{f}(1 - \bar{y}_{ij}) \leq w_{ij}$. Ahora escribimos el problema (RND) dónde podemos evaluar el regret o la pérdida de oportunidad.

$$\begin{aligned}
R^* &:= \text{mín} \quad R(x, y) \\
&\text{sueto a:} \\
&\sum_{(i,j) \in A} x_{ijp} - \sum_{(k,i) \in A} x_{kip} = b_{ip}, \quad i \in N, p \in P, \\
&x_{ijp} \leq y_{ij}, \quad (i, j) \in A, p \in P, \\
&y_{ij} \in \{0, 1\}, \quad (i, j) \in A, \\
&x_{ijp} \in \{0, 1\}, \quad (i, j) \in A, p \in P.
\end{aligned} \tag{RND}$$

Este cambio en la función objetivo del problema (DUND) supone un gran inconveniente para su resolución pues, en general, $R(x, y)$ es una función no lineal en la cual incluso la evaluación de su valor requiere la resolución del problema entero mixto de separación (SP).

Nuestro objetivo va a ser encontrar un diseño robusto resolviendo el problema (RND) mediante un algoritmo de descomposición de Benders. Esta metodología comienza con un diseño dado (x^0, y^0) para el que resolvemos el correspondiente problema de separación que nos dará una cota superior del valor óptimo del problema (RND) y una de sus soluciones óptimas (\bar{x}^0, \bar{y}^0) define el peor adversario bajo el escenario más desfavorable para (x, y) . El diseño de red que se encuentra en cada iteración de dicho método es una solución del problema (RND) para el subconjunto de escenarios de costes generados al resolver el problema de separación (SP) en las iteraciones previas. Definimos el siguiente problema maestro para una iteración n que denotaremos (MP) y que determinará una cota inferior del valor óptimo de (RND).

$$r^* := \text{mín } r$$

sujeto a:

$$\begin{aligned} \sum_{(i,j) \in A} f_{ij}^q y_{ij} + \sum_{(i,j) \in A, p \in P} c_{ij}^q x_{ijp} - D(\bar{x}^q, \bar{y}^q, s^q) &\leq r, q = 0, \dots, n-1, \\ \sum_{(i,j) \in A} x_{ijp} - \sum_{(k,i) \in A} x_{kip} &= b_{ip}, i \in N, p \in P, \\ x_{ijp} &\leq y_{ij}, (i,j) \in A, p \in P, \\ y_{ij} &\in \{0, 1\}, (i,j) \in A, \\ x_{ijp} &\in \{0, 1\}, (i,j) \in A, p \in P, \end{aligned} \tag{MP}$$

La resolución de los problemas de separación y maestro en cada iteración nos permite obtener simultáneamente una cota superior e inferior del valor óptimo del problema (*RND*). El algoritmo de descomposición de Benders se describe a continuación y veremos que acaba en un número finito de iteraciones.

Algorithm 1 Algoritmo de descomposición de Benders

- 1: Inicialización: sea $(x^\epsilon, y^\epsilon) := (x^0, y^0)$ un diseño inicial factible.
 - 2: Resolver el problema de separación para obtener $R(x^0, y^0)$.
 - 3: Sea $s^0 := (c^0, f^0)$ el peor de los escenarios posibles para (x^0, y^0) .
 - 4: Inicializar la lista $L := \{(c^0, f^0)\}$.
 - 5: Sea $\beta := R(x^0, y^0)$ una cota superior del valor óptimo del problema (*RND*).
 - 6: Ir a la iteración $n := 1$.
 - 7: Iteraciones $n = 1, 2, \dots$
 - 8: Resolver el problema maestro (*MP*) y sea (x^n, y^n) una de las soluciones óptimas.
 - 9: Sea $\alpha := r^*$ el valor óptimo del problema (*MP*).
 - 10: Si $\beta - \alpha \leq \epsilon$ (o $(\beta - \alpha)/\alpha \leq \epsilon$) entonces (x^ϵ, y^ϵ) **es una ϵ -aproximación del problema (*RND*)**.
 - 11: En caso contrario, resolver el problema de separación para obtener el valor $R(x^n, y^n)$.
 - 12: Sea $s^n := (c^n, f^n)$ el peor de los escenarios posibles de costes para (x^n, y^n) y (\bar{x}^n, \bar{y}^n) un adversario bajo el peor de los escenarios posibles.
 - 13: Actualizar la lista $L := L \cup \{(c^n, f^n)\}$.
 - 14: Si $R(x^n, y^n) < \beta$ actualizar la cota superior $\beta := R(x^n, y^n)$ y $(x^\epsilon, y^\epsilon) := (x^n, y^n)$
 - 15: Ir a la iteración $n := n + 1$.
-

Proposición 2.1. *Si $\epsilon=0$, el algoritmo anterior termina en un número finito de iteraciones con una solución óptima del problema (*RND*).*

Demostración. Observemos en primer lugar que podemos suponer que la lista L de

escenarios generados al resolver el problema de separación, está formado por vértices de S . Esto es cierto puesto que $R(x, y)$ es una función lineal convexa a trozos que se debe maximizar sobre un poliedro acotado. Para probar la finitud del algoritmo de Descomposición de Benders veremos que el cardinal de la lista L aumenta en una unidad en cada iteración. En caso contrario, si la cardinalidad de L permanece invariante a lo largo del algoritmo entonces dicho escenario habría sido generado en una iteración anterior pues:

Por (1)

$$R(x^n, y^n) = R(x^n, y^n, s^n) = \max_{q=0,1,\dots,n-1} \{R(x^n, y^n, s^q)\}$$

Por el algoritmo de descomposición de Benders:

$$\max_{q=0,1,\dots,n-1} \{R(x^n, y^n, s^q)\} = \alpha$$

Cómo el conjunto $q = 0, 1, \dots, n-1$ está incluido en el conjunto de todos los posibles escenarios entonces se cumple la igualdad siguiente:

$$\max_{q=0,1,\dots,n-1} R(x, y, s^q) \leq \max_{s \in S} R(x, y, s)$$

Además cómo esta desigualdad se cumple $\forall x, y$ podemos tomar mínimo.

$$\min_{x,y} \max_{q=0,1,\dots,n-1} R(x, y, s^q) \leq \min_{x,y} \max_{s \in S} R(x, y, s)$$

Teniendo en cuenta que $\min_{x,y} \max_{q=0,1,\dots,n-1} R(x, y, s^q) = R(x^n, y^n)$ y $R^* = \min_{x,y} \max_{s \in S} R(x, y, s)$ con R^* es el valor óptimo del problema (RND), tenemos:

$$\max_{q=0,1,\dots,n-1} \{R(x^n, y^n, s^q)\} \leq R^*$$

Cómo (x^n, y^n) es la solución óptima del problema (MP) y R^* es el mínimo de $R(x, y)$,

en particular, menor o igual que $R(x^n, y^n)$, por tanto:

$$R(x^n, y^n) = \max_{q=0,1,\dots,n-1} \{R(x^n, y^n, s^q)\} \leq R^* \leq R(x^n, y^n)$$

Finalmente

$$R(x^n, y^n) \leq R(x^n, y^n)$$

Por tanto debe darse la igualdad en toda la cadena de desigualdades y el algoritmo terminará en la siguiente iteración con la solución óptima (x^n, y^n) .

□

2.1.5. Ejemplo (DUND) (*continuación*)

En la primera parte de este ejemplo, asumíamos que los costos fijos y unitarios de los arcos de transporte eran conocidos. Consideraremos ahora que algunos de estos costes son inciertos. El origen de la incertidumbre puede ser diverso. Pensemos, por ejemplo, que estamos diseñando una red de transporte en la que tenemos que cavar un túnel a través de una montaña. Su coste de construcción puede ser muy variable dependiendo, por ejemplo, de si tenemos dificultades en las tareas de excavación o no. En otros proyectos, la fuente de la incertidumbre podría ser la falta de experiencia realizando algunas de las tareas de construcción correspondientes. Por ejemplo, en el *Haramain Ferrocarril* de alta velocidad recientemente abierto al público en Arabia Saudita, que conecta ciudades de Medina y La Meca, la poca experiencia en la construcción de raíles sobre arena fue una de las principales fuentes de incertidumbre sobre los costes de construcción y mantenimiento.

Para ilustrar tales incertidumbres en nuestro ejemplo de juguete, identificaremos con un rectángulo obstáculo, un conjunto de arcos potenciales cuyos costes fijos pueden variar en una cantidad incierta. En la figura 4, esta zona incierta ha sido representada por un rectángulo gris dónde los arcos dentro del rectángulo son los afectados por la incertidumbre en sus costes. Suponemos que los costes fijos para todos esos arcos ahora pueden variar, aumentando o disminuyendo su valor de 1 unidad en un máximo de 0,6

unidades, es decir, pueden tomar cualquier valor en el intervalo $[0,4,1,6]$.

La figura 5 muestra el diseño de la red que minimiza el máximo arrepentimiento bajo esta nueva estructura de costes resolviendo el problema (RND) mediante el algoritmo de Descomposición de Benders. Podemos observar que la estructura de la red ha cambiado significativamente para evitar la construcción de los arcos potenciales con costes fijos inciertos. Entonces, podemos observar que incorporar costes inciertos al diseño de una red puede cambiar la estructura de la solución tratando de esquivar, si es posible, la construcción de arcos potenciales con costes inciertos. Este cambio en el diseño de la red parece ser una respuesta razonable (robusta) a la nueva estructura de costes impredecible y muestra la sensibilidad de nuestro modelo a los costes inciertos.

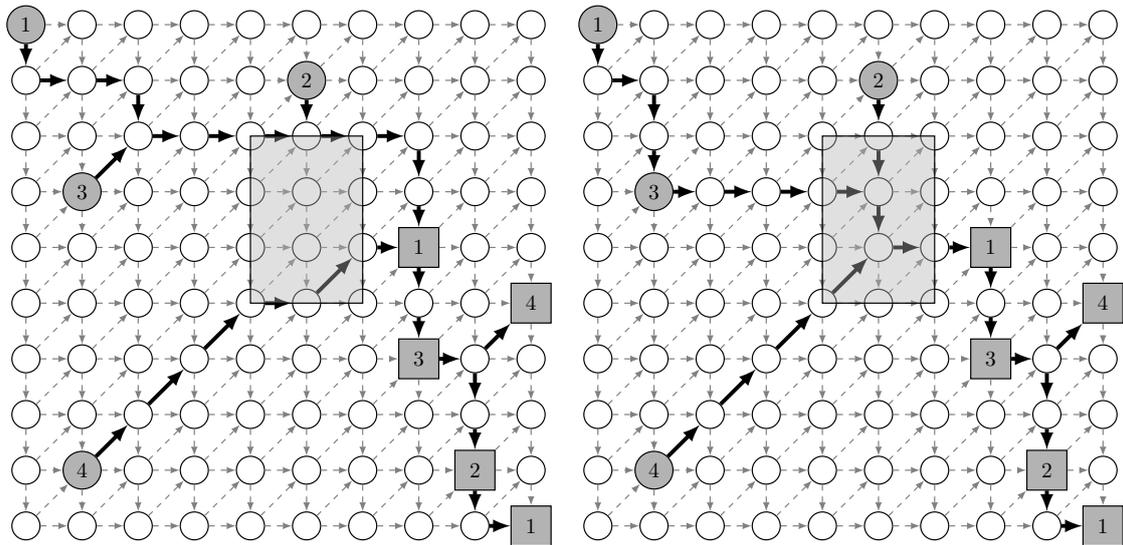


Figura 5: Diseño óptimo asumiendo costes inciertos y pares vinculados.

Figura 6: Diseño óptimo sin incertidumbre.

2.2. Diseño robusto bajo costes inciertos y pares vinculados.

En esta sección la lista de pares enlazados se conoce parcialmente modelando situaciones en las que los proveedores y clientes son parcialmente conocidos. La lista de pares P anteriores se convierte ahora en la lista de pares que efectivamente deberían de estar conectados bajo cada escenario S . En este caso existirán pares de nodos que finalmente no se conectan y es por ello que no tendrán coste de transporte pero sí de mantenimiento, por ejemplo, si construimos una carretera que resulta ser inutilizada únicamente se producirían costes para mantenerla. Los escenarios posibles según la lista de pares vinculados pueden escribirse como restricciones lineales en un conjunto de variables binarias:

$$\sum_{p \in P} e_{pr} u_p \leq g_r, \quad r \in R, u_p \in \{0, 1\}, \forall p \in P$$

dónde u_p es una variable binaria que toma el valor 1 si el par $p \in P$ está conectado y 0 en caso contrario. Ahora, e_{pr} es una variable técnica, por ejemplo si suponemos $e_{pr} = 1$ y $g_r = 7$ entonces eso significará que cómo mucho 7 pares de nodos pueden vincularse. R identifica los requisitos (restricciones) sobre los pares que deben vincularse.

2.2.1. Formulación del problema de separación y maestro.

Cómo en el modelo anterior, la función del problema de separación será la de determinar el peor escenario posible para un diseño candidato que se parametriza mediante las variables (x, y) . En el nuevo modelo dicho escenario se compone no sólo de una realización de costes desfavorable sino además de los pares vinculados que deberá unir el diseño candidato. Para contemplar esta nueva situación añadimos un conjunto de restricciones y de nuevas variables a la formulación del problema de separación que ahora denotaremos por (*SPUL*).

$$R(x, y) := \text{máx} \quad \sum_{(i,j) \in A} f_{ij} y_{ij} + \sum_{(i,j) \in A, p \in P} \xi_{ijp} x_{ijp} \\ - \sum_{(i,j) \in A} w_{ij} - \sum_{(i,j) \in A, p \in P} z_{ijp}$$

sujeto a:

$$\sum_{(i,j) \in A} \bar{x}_{ijp} - \sum_{(k,i) \in A} \bar{x}_{kip} = b_{ip}, \quad i \in N, p \in P,$$

$$\bar{x}_{ijp} \leq \bar{y}_{ij}, \quad (i, j) \in A, p \in P,$$

$$\sum_{(i,j) \in A} (a_{ijk}^f f_{ij} + a_{ijk}^c c_{ij}) \leq d_k, \quad k \in K,$$

$$\sum_{p \in P} e_{pr} \bar{u}_p \leq g_r, \quad r \in R,$$

$$c_{ij} - \bar{c}(2 - \bar{u}_p - \bar{x}_{ijp}) \leq z_{ijp}, \quad (i, j) \in A, p \in P,$$

$$0 \leq z_{ijp}, \quad (i, j) \in A, p \in P,$$

$$\xi_{ijp} \leq c_{ij}, \quad (i, j) \in A, p \in P,$$

$$0 \leq \xi_{ijp} \leq \bar{c} \bar{u}_p, \quad (i, j) \in A, p \in P,$$

$$f_{ij} - \bar{f}(1 - \bar{y}_{ij}) \leq w_{ij}, \quad (i, j) \in A,$$

$$0 \leq w_{ij}, \quad (i, j) \in A,$$

$$\bar{y}_{ij} \in \{0, 1\}, \quad (i, j) \in A,$$

$$\bar{x}_{ijp} \in \{0, 1\}, \quad (i, j) \in A, p \in P,$$

$$\bar{u}_p \in \{0, 1\}, \quad p \in P,$$

$$f_{ij} \geq 0, \quad (i, j) \in A,$$

$$c_{ij} \geq 0, \quad (i, j) \in A.$$

(SPUL)

Añadimos la restricción del número máximo de nodos a vincular que explicamos anteriormente. La restricción $c_{ij} - \bar{c}(2 - \bar{u}_p - \bar{x}_{ijp}) \leq z_{ijp}$, $(i, j) \in A, p \in P$ se explica cómo en el problema de separación. Si no conectamos el par $p \in P$ entonces $\bar{u}_p = 0$ y $\bar{x}_{ijp} = 0$. Por tanto la restricción queda como sigue: $c_{ij} - 2\bar{c} \leq z_{ijp}$ con \bar{c} suficientemente

grande. Cómo también tenemos la restricción $z_{ijp} \geq 0$ al unir las dos y teniendo en cuenta que estamos maximizando entonces $z_{ijp} = 0$. Al igual ocurre cuando el par $p \in P$ está conectado, entonces $\bar{u}_p = \bar{x}_{ijp} = 1$. Al estar maximizando buscamos la igualdad de $c_{ij} = z_{ijp}$ teniendo en cuenta que \bar{c} es suficientemente grande. En la restricción $0 \leq \xi_{ijp} \leq \bar{c} \bar{u}_p$ cómo tenemos $\xi_{ijp} \leq c_{ijp} \leq \bar{c}$ para asegurarnos de que sólo debe agregarse la conexión si $\bar{u}_p = 1$. Entonces ocurre que $\xi_{ijp} \leq \bar{c} \bar{u}_p$ en caso de que no se enlacen dichos nodos los ξ_{ijp} son cero y por tanto el término de la función objetivo dónde aparece no se añadiría. Es muy importante para esta formulación las variables continuas ξ_{ijp} pues no tendría sentido incluir un término en la función objetivo con el coste de un arco de transporte que no está construido. La formulación (*SPUL*) tendrá como solución un escenario de costes (f, c) bajo un diseño (N, A_y) que mostrará su faceta más vulnerable frente a su peor adversario bajo el escenario más desfavorable posible.

Notemos que el bloque de restricciones

$$\sum_{(i,j) \in A} \bar{x}_{ijp} - \sum_{(k,i) \in A} \bar{x}_{kip} = b_{ip}, \quad i \in N, p \in P$$

es necesaria pues sino estaríamos incluyendo una asimetría en el modelo. El adversario bajo el peor de los escenarios posibles no tendría porqué incluir en su solución la conexión de todos los pares o incluso la construcción de todos los arcos que aseguren dicha conexión. Por tanto, la asimetría viene del hecho de comparar un posible diseño (x, y) con el adversario bajo el peor de los escenarios posibles que no obliga a tener las mismas conexiones entre nodos. Sin embargo, los costes de transporte entre nodos sólo se consideran entre los pares que aparecen en el peor de los escenarios posibles para poder comparar un diseño dado con este último.

Por lo tanto, para cualquier solución dada (x, y) resolvemos el problema de separación (*SPUL*) para encontrar el adversario bajo el peor de los escenarios posibles y con ello el conjunto de pares que deben estar vinculados junto a sus costes. El problema maestro que usaremos es muy parecido al del modelo anterior (MP) cómo se puede ver a continuación:

mín r

sujeto a:

$$\begin{aligned}
& \sum_{(i,j) \in A} f_{ij}^q y_{ij} + \sum_{(i,j) \in A, p \in P} c_{ij}^q \bar{u}_p^q x_{ijp} - D(\bar{x}^q, \bar{y}^q, s^q) \leq r, q = 0, 1, \dots, n-1, \\
& \sum_{(i,j) \in A} x_{ijp} - \sum_{(k,i) \in A} x_{kip} = b_{ip}, i \in N, p \in P, \\
& x_{ijp} \leq y_{ij}, (i, j) \in A, p \in P, \\
& y_{ij} \in \{0, 1\}, (i, j) \in A, \\
& x_{ijp} \in \{0, 1\}, (i, j) \in A, p \in P,
\end{aligned} \tag{MPUL}$$

El valor óptimo del problema (*MPUL*) es una cota inferior del valor óptimo del problema (*RND*) bajo un marco de incertidumbre en los costes y conexión de nodos. Además el sumando $\sum_{(i,j) \in A, p \in P} c_{ij}^q \bar{u}_p^q x_{ijp}$ sólo se incluye si el par $p \in P$ está conectado bajo el escenario s^q . El primer bloque de restricciones compara el coste total de construcción y conexión entre nodos en el diseño considerado bajo el escenario de costes y pares vinculados. El segundo bloque son restricciones de flujo que obliga a conectar cada par de nodos potencialmente vinculados.

El algoritmo 2 incluye las adaptaciones necesarias en el esquema de Descomposición de Benders para resolver el nuevo modelo de optimización. Su convergencia sigue estando garantizada ya que se mantiene la finitud del conjunto de escenarios que serán generados mediante el problema de separación.

Proposición 2.2. *Si $\epsilon = 0$ este último algoritmo acaba en un número finito de iteraciones con una solución óptima del problema (*RND*) bajo una estructura incierta de pares enlazados.*

Demostración. La formulación (*SPUL*) nos garantiza que se alcanza la solución óptima de (*RND*) ya que se puede encontrar un punto extremo factible en el poliedro acotado de escenarios de costes. Como antes, asumimos que la lista L es finita pues el conjunto factible de subconjuntos de pares enlazados lo es. Para probar la finitud del algoritmo mostraremos que la cardinalidad de la lista L aumenta en cada iteración. El conjunto

Algorithm 2 Algoritmo de descomposición de Benders 2

- 1: Sea $(x^\epsilon, y^\epsilon) := (x^0, y^0)$ un diseño factible.
 - 2: Resolver el problema de separación (*SPUL*) para obtener $R(x^0, y^0)$.
 - 3: Sea $s^0 := (c^0, f^0, u^0)$ el adversario bajo el peor de los escenarios posibles para (x^0, y^0) .
 - 4: Inicializar la lista $L := \{(c^0, f^0, u^0)\}$.
 - 5: Sea $\beta := R(x^0, y^0)$ una cota superior del valor óptimo del problema (*RND*) con incertidumbre sobre los nodos que deberán ser conectados.
 - 6: Ir a la iteración $n := 1$.
 - 7: Iteraciones $n = 1, 2, \dots$
 - 8: Resolver el problema maestro (*MPUL*) y sea (x^n, y^n) una de sus soluciones óptimas.
 - 9: Sea $\alpha := r^*$ un valor óptimo del problema (*MPUL*).
 - 10: Si $\beta - \alpha \leq \epsilon$ (o $(\beta - \alpha)/\alpha \leq \epsilon$)
 - 11: Entonces (x^ϵ, y^ϵ) es una ϵ -aproximación del problema (*RND*).
 - 12: Resolver el problema (*SPUL*) para obtener $R(x^n, y^n)$.
 - 13: Sea $s^n := (c^n, f^n, u^n)$ sea el peor escenario posible para (x^n, y^n) y $(\bar{x}^n, \bar{y}^n, \bar{u}^n)$ el peor adversario bajo el peor de los escenarios posibles.
 - 14: Actualizar la lista $L := L \cup \{(c^n, f^n, u^n)\}$.
 - 15: Si $R(x^n, y^n) < \beta$
 - 16: Actualizar la cota superior $\beta := R(x^n, y^n)$ y $(x^\epsilon, y^\epsilon) := (x^n, y^n)$
 - 17: Ir a la iteración $n := n + 1$.
-

de posibles escenarios se puede considerar finito y por tanto el número de iteraciones también lo será. Si la lista L en una de las iteraciones permaneciera sin cambios entonces dicho escenario ya habría sido generado en una iteración anterior. Por tanto:

En este caso $s^n := (c^n, f^n, u^n)$.

$$R(x^n, y^n) = R(x^n, y^n, s^n) = \max_{q=0,1,\dots,n-1} \{R(x^n, y^n, s^q)\}$$

Por el algoritmo de descomposición de Benders modificado:

$$\max_{q=0,1,\dots,n-1} \{R(x^n, y^n, s^q)\} = \alpha$$

Al igual que en la anterior proposición, tomando máximo y después mínimo teniendo en cuenta la inclusión de $q = 0, 1, \dots, n - 1$ en el conjunto de todos los posibles escenarios tenemos:

$$\max_{q=0,1,\dots,n-1} \{R(x^n, y^n, s^q)\} \leq R^*$$

dónde R^* es el valor óptimo del problema (RND). Como antes:

$$R(x^n, y^n) = \max_{q=0,1,\dots,n-1} \{R(x^n, y^n, s^q)\} \leq R^* \leq R(x^n, y^n)$$

Finalmente

$$R(x^n, y^n) \leq R(x^n, y^n)$$

Por tanto debe darse la igualdad en toda la cadena de desigualdades y el algoritmo terminará en la siguiente etapa con la solución óptima (x^n, y^n) .

□

2.2.2. Ejemplo (DUND) (*continuación*)

Para concluir con el ejemplo desarrollado a lo largo de estas secciones consideramos incertidumbre en la demanda que podría modelar situaciones reales en las que el tráfico o el flujo de mercancías entre los nodos origen-destino sólo se conoce parcialmente. Un ejemplo de esta situación se puede encontrar en el tráfico existente en las redes de comunicación. En la mayoría de los casos, no existe información precisa sobre el tráfico esperado entre las terminales a conectar, sin embargo, se podría disponer de una colección de posibles matrices de tráfico. Con el objetivo de representar esta situación, supondremos que el conjunto de demandas posibles (para cada uno de los cinco pares que podrían vincularse bajo un posible escenario) es conocido. Estos conjuntos van a ser respectivamente, $\{0, 1, 2, 3\}$, $\{0, 1, 2\}$, $\{0, 1\}$, $\{0, 1\}$ y $\{0, 1, 2, 3\}$ para el conjunto ordenado de pares de P . Los valores considerados para las demandas podrían representar el valor de tráfico habitual como 1, el doble del valor de tráfico habitual como 2 y así el valor 0 representa la ausencia de tráfico entre el par de nodos.

El diseño óptimo para el ejemplo asumiendo costes inciertos y pares vinculados representado en La figura 7 no es óptima en esta nueva variante del problema (RND), pues logra un máximo arrepentimiento, bajo estas nuevas condiciones, del 58.93% superior a su valor óptimo. Esto demuestra, una vez más, la sensibilidad del diseño óptimo bajo variación de las distintas fuentes de incertidumbre.

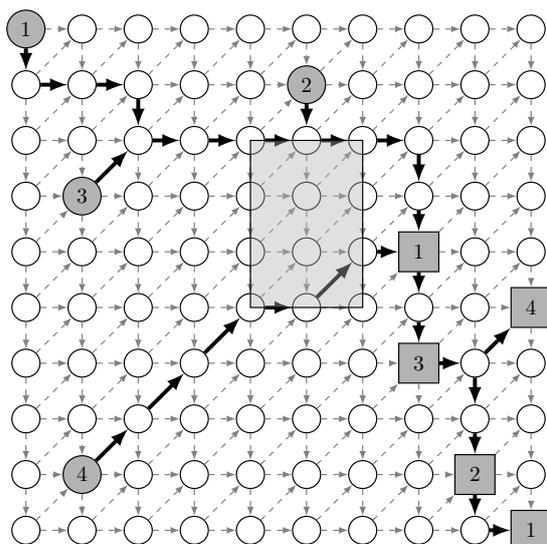


Figura 7: Diseño óptimo asumiendo costes y pares vinculados inciertos.

2.3. Implementación algorítmica mediante AMPL

En esta sección presentaremos el código AMPL utilizado para implementar el algoritmo de Descomposición de Benders. Recordemos que dicho algoritmo busca un diseño robusto de la red de acuerdo al modelo (RND). Cómo ya sabemos, comienza con un diseño dado (x^0, y^0) y resuelve el correspondiente problema de separación que proporciona una cota superior del valor óptimo del problema (RND) y una de sus soluciones óptimas definirá el adversario bajo el peor de los escenarios posibles (\bar{x}^0, \bar{y}^0) . En cada iteración obtenemos un diseño candidato para el problema (RND) que es óptimo para el subconjunto de escenarios considerado. Para poder implementarlo haremos uso del problema maestro (MP) y el problema de separación (SP), que es la formulación resultante de linealizar el problema (1) bajo nuestro poliedro S .

2.3.1. Código AMPL del problema maestro (MP) y del de separación (SP)

Comenzamos definiendo los parámetros que vamos a usar en el problema.

- n define el número de nodos que tenemos, según la estructura de red representada en la figura 3, la red potencial va a tener $n \times n$ nodos.

- el parámetro *pp* será el número de pares a vincular, ya que en esta formulación los supondremos conocidos.

```
param n;
```

```
param pp;
```

Definimos I como un conjunto de índices que abarca desde 1 hasta n y sea P el conjunto de índices de los pares que deben vincularse. Notemos que para poder definir dichos conjuntos es necesario haber definido los parámetros anteriores.

```
set I:=1..n;
```

```
set P:=1..pp;
```

A continuación vamos a construir los arcos potenciales de la red en formato malla. Teniendo en cuenta la estructura particular del grafo de conexiones potenciales, parece adecuado denotar a los nodos por pares de índices que identifican su fila y columna dentro de la malla, (véase la figura 3). Recordemos que dichos arcos son potenciales porque en realidad no están construidos y según nuestra solución decidiremos si serán incluidos en el diseño de la red.

El comando *setof* es un operador de construcción de conjuntos, puede ser una expresión o una lista de expresiones separadas por comas entre paréntesis. El conjunto final será el resultado de iterar sobre la expresión de indexación, en este caso i desde 1 hasta n y j desde 1 hasta $n - 1$. La dimensión de la expresión resultante es el número de componentes dentro del comando *setof*. Los primeros arcos construidos son los horizontales ya que teniendo en cuenta que cada fila y cada columna consta de n nodos, comenzamos en el nodo inicial (i, j) con $i = j = 1$ y sustituyendo tenemos que el primer arco iría al $(1, 2)$ es decir el recorrido de dicho arco sería del $(1, 1)$ al $(1, 2)$ por tanto dicho arco potencial debe ser horizontal. Los siguientes arcos construidos son los verticales razonando del mismo modo. Empezamos en el nodo inicial $(1, 1)$, tomamos $i = j = 1$ y sustituyendo en la segunda expresión tenemos que el primer arco iría al $(1, n + 1)$, es decir está en posición vertical ya que en cada fila tenemos n nodos y el $n + 1$ se correspondería con el primer nodo de la segunda fila. Por último, la expresión final del código se corresponden con

los arcos oblicuos ya que del arco inicial sustituyendo $i = j = 1$ iríamos al $(n + 1, n + 2)$ que con n nodos en la primera fila correspondería al segundo elemento de la segunda fila. Para definir todos los arcos a la vez hacemos uso del comando *union*, que cómo su nombre indica se usa para la unión de conjuntos.

```
set arcos := setof {i in 1..n,j in 1..(n-1)} (n*(i-1)+j,n*(i-1)+j+1)
union setof {i in 1..(n-1),j in 1..n} (n*(i-1)+j,n*i+j)
union setof {i in 2..n,j in 1..(n-1)} (n*(i-1)+j,n*(i-2)+j+1);
```

Una vez definido los distintos tipos de arcos continuamos con los parámetros siguientes:

```
param cmenos{arcos};
param cmas{arcos};
param fmenos{arcos};
param fmas{arcos};
param od{1..2,P};
```

En nuestro modelo consideramos en general el siguiente poliedro $\sum_{(i,j) \in A} (a_{ijk}^f f_{ij} + a_{ijk}^c c_{ij}) \leq d_k, k \in K$, pero por simplicidad en la implementación, vamos a considerar como conjunto de escenarios posibles $s \in S$ un hipercubo, es decir, consideramos los costes de transporte y mantenimiento de arcos c_{ij} y f_{ij} acotados inferior y superiormente. En este caso, las cotas inferiores y superiores del coste c_{ij} serían los parámetros *cmenos* y *cmas* respectivamente. Análogamente, el coste f_{ij} está acotado por *fmenos* y *fmas* respectivamente. El último parámetro corresponde al origen-destino, es una variable con dos componentes correspondientes al número de pares a vincular, por ejemplo si tenemos que transportar unidades del primer nodo hasta el noveno entonces $od(1, 1) = 1$ y $od(2, 1) = 9$. En el ejemplo consideramos sólo un par de nodos a vincular.

Por tanto para poder implantar el algoritmo de Descomposición de Benders 1 necesitamos el problema maestro y de separación, entonces empezamos implementando este último como sigue:

```

param xx{arcos,P}>=0;
param yy{arcos}>=0;

```

El parámetro xx representa el flujo x_{ijp} , cómo vemos depende de los arcos potenciales y del índice de los pares que deben vincularse. El siguiente parámetro yy debe ser positivo y depende únicamente de los arcos potenciales anteriores, representa los valores y_{ij} que definen un diseño de la red.

```

var xbar{arcos,P} binary;
var ybar{arcos} binary;

```

El código define las variables de decisión para el solver usado por AMPL y hacen referencia respectivamente a los \bar{x}_{ijp} y \bar{y}_{ij} del problema de optimización (1). Hacemos uso de las barras debido a que x e y se usan en el *input* de $R(x, y)$. Además debemos exigir que las variables sean binarias puesto que el problema de separación no es unimodular.

Definimos a continuación la variable c_{ijp} que es el coste de transporte a través del arco (i, j) bajo el escenario $s \in S$ y el coste de construcción del arco potencial (i, j) es f_{ij} , que en la implementación se corresponden con $cbar$ y $fbar$. Una vez obtenido el problema de optimización (1) para poder linealizar los términos cuadráticos que aparecían en la función objetivo definíamos $c_{ij}\bar{x}_{ijp} = z_{ijp}$ y $f_{ij}\bar{y}_{ij} = w_{ij}$ que se corresponderían en este caso con las variables $zbar$ y $wbar$, ambas positivas.

```

var cbar{arcos}>=0;
var fbar{arcos}>=0;
var zbar{arcos,P}>=0;
var wbar{arcos}>=0;

```

Vamos a definir la función objetivo del problema de separación dónde los sumandos se encuentran escritos en el mismo orden que aparecen implementados para mejor comprensión.

$$R(x, y) := \max \sum_{(i,j) \in A} f_{ij} y_{ij} + \sum_{(i,j) \in A, p \in P} c_{ij} x_{ijp} - \sum_{(i,j) \in A} w_{ij} - \sum_{(i,j) \in A, p \in P} z_{ijp}$$

```

maximize R(x,y):
sum {(i,j) in arcos} fbar[i,j]*yy[i,j]+
sum {(i,j) in arcos,p in P} cbar[i,j]*xx[i,j,p]-
sum {(i,j) in arcos} wbar[i,j]-
sum {(i,j) in arcos,p in P} zbar[i,j,p];

```

Pasamos ahora a las siguientes restricciones del problema de separación, que están escritas en diez bloques (denotados por B1bar, ..., B10bar) para aprovechar la estructura formal del problema de optimización.

```

subject to B1bar {i in {1..n^2}, p in P: od[1,p]==i}:
sum {(i,j) in arcos} xbar[i,j,p]-sum {(k,i) in arcos} xbar[k,i,p]= 1;
subject to B2bar {i in {1..n^2}, p in P: od[2,p]==i}:
sum {(i,j) in arcos} xbar[i,j,p]-sum {(k,i) in arcos} xbar[k,i,p]= -1;
subject to B3bar {i in {1..n^2}, p in P: od[1,p]<>i and od[2,p]<>i}:
sum {(i,j) in arcos} xbar[i,j,p]-sum {(k,i) in arcos} xbar[k,i,p]= 0;

```

Estos tres bloques de restricciones se corresponde con la siguiente formulación:

$$\sum_{(i,j) \in A} \bar{x}_{ijp} - \sum_{(k,i) \in A} \bar{x}_{kip} = b_{ip}, \quad i \in N, p \in P$$

El bloque B1bar se usa en el código para identificar las restricciones de balance de flujo correspondientes a los nodos origen del conjunto de pares que deben ser conectados, en este caso $b_{ip} = 1$. El bloque B2bar identifica las restricciones de flujo de los nodos destino, $b_{ip} = -1$. El bloque B3bar representan las restricciones de los nodos intermedios (transbordo) del flujo que se emitía entre nodos de uno de los pares que deben ser conectados, en este caso $b_{ip} = 0$.

Si no construimos el arco que une el nodo i con j entonces no podremos transportar las unidades a través de dicho arco, entonces añadimos la restricción $\bar{x}_{ijp} \leq \bar{y}_{ij}$, $(i, j) \in A, p \in P$ que se corresponde con:

```

subject to B4bar {(i,j) in arcos,p in P}:

```

xbar[i,j,p]<=ybar[i,j];

El siguiente bloque corresponde a la implementación de las siguientes restricciones:

$$c_{ij} - \bar{c}(1 - \bar{x}_{ijp}) \leq z_{ijp}, \quad (i, j) \in A, p \in P$$

$$f_{ij} - \bar{f}(1 - \bar{y}_{ij}) \leq w_{ij}, \quad (i, j) \in A$$

subject to B5bar {(i,j) in arcos,p in P}:

cbar[i,j]-cmas[i,j]*(1-xbar[i,j,p])<=zbar[i,j,p];

subject to B6bar {(i,j) in arcos}:

fbar[i,j]-fmas[i,j]*(1-ybar[i,j])<=wbar[i,j];

Para el siguiente bloque de restricciones recordamos el poliedro en general que estábamos usando: $\sum_{(i,j) \in A} (a_{ijk}^f f_{ij} + a_{ijk}^c c_{ij}) \leq d_k, k \in K$. Entonces hemos dicho que vamos a considerar como escenario S un hipercubo y que por tanto los costes c_{ij} y f_{ij} estarán acotadas inferior y superiormente y por tanto las siguientes restricciones implementadas corresponden con: $c_{ij} \leq c_{ij}^+, f_{ij} \leq f_{ij}^+, c_{ij} \geq c_{ij}^-, f_{ij} \geq f_{ij}^-$ respectivamente.

subject to B7bar {(i,j) in arcos}:

cbar[i,j]<=cmas[i,j];

subject to B8bar {(i,j) in arcos}:

fbar[i,j]<=fmas[i,j];

subject to B9bar {(i,j) in arcos}:

cbar[i,j]>=cmenos[i,j];

subject to B10bar {(i,j) in arcos}:

fbar[i,j]>=fmenos[i,j];

A continuación describimos la implementación del *AMPL* del problema maestro, para ello definimos los siguientes parámetros:

param nCUT >= 0 integer;

param c{arcos,1..nCUT};

```

param f{arcos,1..nCUT};
param D{1..nCUT};

```

El primero de ellos corresponde con el número de cortes que se incluyen en el problema maestro (MP). Los dos siguientes son los escenarios de coste dónde se alcanza el regret máximo para alguna de las soluciones candidatas que se han generado. El último parámetro corresponde al regret máximo de las soluciones candidatas que se han generado en el algoritmo.

Definimos las mismas variables que en el problema de separación:

- x_{ijp} depende de los arcos y del número de pares a vincular y simboliza una variable de flujo binario que toma el valor uno si se elige dicho arco para conectar al par $p \in P$.
- la variable binaria y_{ij} depende únicamente de los arcos y toma el valor uno si se construye el arco (i, j) .
- La última variable corresponde con el regret máximo que se debe minimizar en el problema maestro.

```

var x{arcos,P} binary;
var y{arcos} binary;
var r;

```

Definimos a continuación la función objetivo del problema maestro que se corresponde con $r^* = \text{mín } r$.

```

minimize MP: r;

```

Igual que en el problema de separación, se han agrupado las restricciones del problema (MP) en cinco bloques de acuerdo a su estructura formal. Los tres primeros bloques de restricciones corresponden con la ecuación de balance de flujo $\sum_{(i,j) \in A} x_{ijp} - \sum_{(k,i) \in A} x_{kip} = b_{ip}, i \in N, p \in P$. Al igual que antes, si el origen del par P es el nodo i entonces $b_{ip} = 1$ y una unidad de mercancía sale de i . Si dicho nodo es el destino entonces $b_{ip} = -1$ y una unidad se queda en el nodo i , en caso contrario $b_{ip} = 0$.

```

subject to B1 {i in {1..n^2}, p in P: od[1,p]==i}:
sum {(i,j) in arcos} x[i,j,p]-sum {(k,i) in arcos} x[k,i,p]= 1;
subject to B2 {i in {1..n^2}, p in P: od[2,p]==i}:
sum {(i,j) in arcos} x[i,j,p]-sum {(k,i) in arcos} x[k,i,p]= -1;
subject to B3 {i in {1..n^2}, p in P: od[1,p]<>i and od[2,p]<>i}:
sum {(i,j) in arcos} x[i,j,p]-sum {(k,i) in arcos} x[k,i,p]= 0;

```

Para evitar transportar la mercancía por un arco que no está construido imponemos $x_{ijp} \leq y_{ij}, (i, j) \in A, p \in P$ que se implementa como sigue:

```

subject to B4 {(i,j) in arcos,p in P}:
x[i,j,p]<=y[i,j];

```

El último bloque de restricciones fuerza a la variable r a tomar el máximo regret en el conjunto de escenarios generados en iteraciones previas del algoritmo, es decir, $\sum_{(i,j) \in A} f_{ij}^q y_{ij} + \sum_{(i,j) \in A, p \in P} c_{ij}^q x_{ijp} - D(\bar{x}^q, \bar{y}^q, s^q) \leq r, q = 0, \dots, n-1$ que se codifica en *AMPL* cómo sigue:

```

subject to CUTS {k in 1..nCUT}:
sum {(i,j) in arcos} f[i,j,k]*y[i,j]+
sum {(i,j) in arcos,p in P} c[i,j,k]*x[i,j,p]-D[k]<=r;

```

2.3.2. Redes obtenidas durante la aplicación del algoritmo

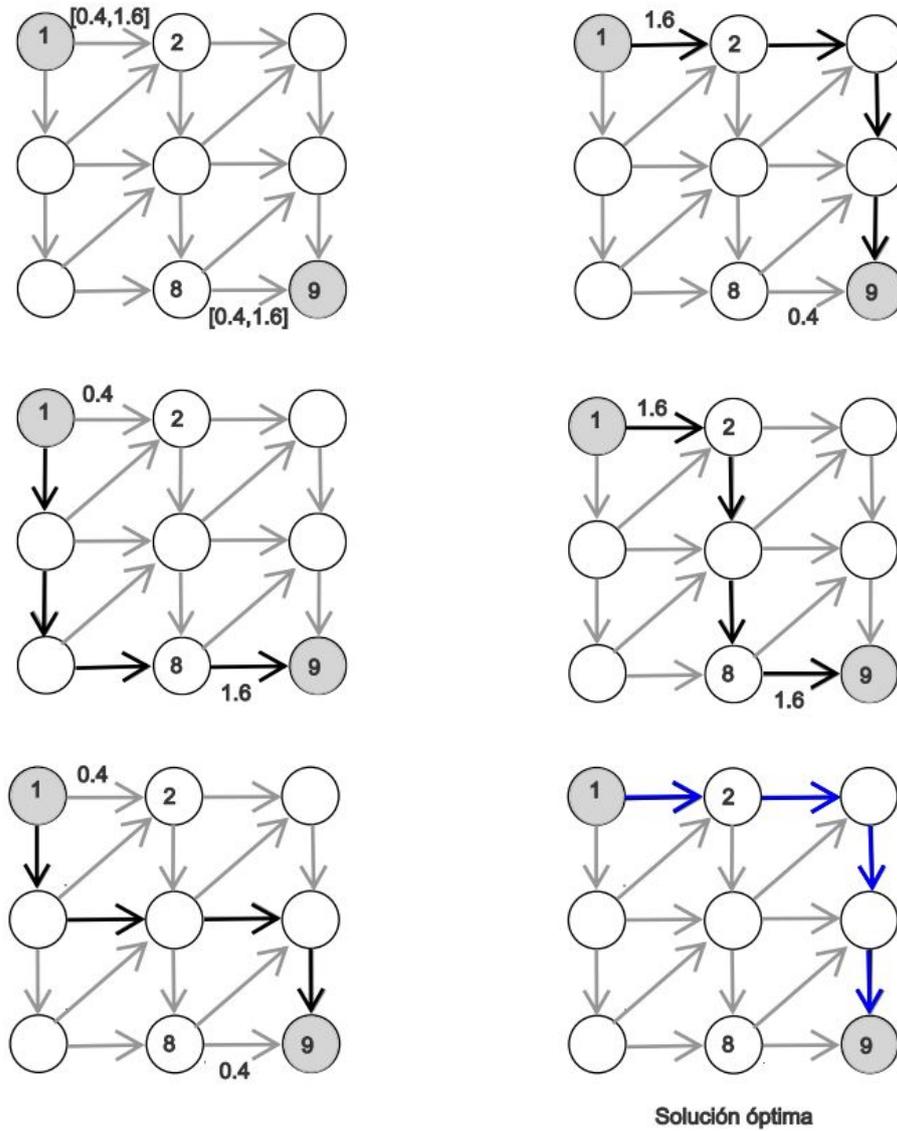


Figura 8: Aplicación del algoritmo.

En este ejemplo consideramos todos los arcos potenciales de una malla 3×3 donde los costes de mantenimiento son conocidos en todos los arcos salvo en dos de ellos y valen una unidad. El arco del nodo 1 al 2 y del 8 al 9 tienen un coste de construcción que varía entre $[0,4, 1,6]$. En la primera red de la figura 8 se presenta consideramos la iteración $n = 0$ con un diseño factible inicial y resolvemos el problema de separación

correspondiente para medir el máximo arrepentimiento y consideramos un escenario de costes para el adversario bajo el peor de los escenarios, en este caso sería considerar el coste del arco del nodo 8 al 9 cómo 0.4, es decir tendría el menor coste de todos los arcos de la malla y por dicho arco no pasa nuestra solución y el otro arco con incertidumbre pasaría a tener un coste de 1.6 unidades. Actualizaremos la cota superior del problema (RND) y resolvemos el problema maestro reiteradamente variando nuestra solución cómo se representa en el resto de mallas hasta llegar a la solución óptima representada en el último dibujo, que tiene un máximo regret de 1.2.

Cómo se puede ver, el diseño robusto de la red evita uno de los arcos con incertidumbre. En el ejemplo considerado, la elección de los intervalos de incertidumbre cómo $[0,4, 1,6]$ hace que existan múltiples soluciones óptimas. Sin embargo, modificando ligeramente dicho intervalo, por ejemplo $[0,6, 1,6]$, el algoritmo 1 detecta cómo diseño óptimo el de la penúltima red representada en la figura 8, es decir, la solución pasa por los nodos 1,4,5,6,9 con un máximo regret de 0.8 y que evita ambos arcos con coste incierto.

2.3.3. Detalle de la salida del algoritmo

Cómo se ilustra en la primera red de la sección anterior, inicializamos el algoritmo con un diseño factible cuya solución parte del nodo 1 y llega al 9 a través del 2,3 y 6. Para poder medir el máximo arrepentimiento de la solución, elegimos como el peor escenario de costes posibles fijando 1.6 el coste del arco que va del nodo 1 al 2 ya que pertenece a la solución y el otro coste de construcción del arco con incertidumbre 0.4, ya que no pasa la solución por él y es el coste más bajo de este arco lo que significa que representaría el escenario más desfavorable para la solución evaluada. Por tanto dicho escenario haría arrepentirnos ya que si hubiésemos sabido dichos costes optaríamos por modificar la solución. Resolvemos el problema de separación correspondiente y obtenemos una cota superior del problema (*RND*) que en este caso es $\beta = 1,2$.

INICIALIZACIÓN n=0, (x⁰,y⁰)

CPLEX 20.1.0.0: optimal integer solution; objective 5.2

0 MIP simplex iterations

0 branch-and-bound nodes

Solución actual $\hat{y}^n =$

$y :=$

1 2 1

1 4 0

2 3 1

2 5 0

3 6 1

4 2 0

4 5 0

4 7 0

5 3 0

5 6 0

5 8 0

6 9 1

7 5 0

7 8 0

8 6 0

8 9 0;

CPLEX 20.1.0.0: optimal integer solution; objective 1.2

0 MIP simplex iterations

0 branch-and-bound nodes

absmipgap = 4.44089e-16, relmipgap = 3.70074e-16

Máximo regret $R(\hat{x}^n, \hat{y}^n) = 1.200000$

Escenario más desfavorable $\hat{s}^n = (\hat{c}^n, \hat{f}^n)$

\bar{f} :=

1 2 1.6

1 4 1

2 3 1

```

2 5 1
3 6 1
4 2 1
4 5 1
4 7 1
5 3 1
5 6 1
5 8 1
6 9 1
7 5 1
7 8 1
8 6 1
8 9 0.4;

```

Cota superior del regret óptimo beta = 1.2

En la siguiente iteración tenemos la solución que parte del nodo 1 y llega al 9 a través de los nodos 4,7 y 8. El peor escenario de costes posibles en este caso sería elegir 1.6 como el coste del arco que va del nodo 8 al 9, ya que dicho arco se encuentra en la solución y el otro arco con incertidumbre por la que no pasa la solución tendría un coste de 0.4 unidades. Resolvemos el problema maestro y obtenemos una cota inferior α nula (aunque *CPLEX* la representa con el valor $-4.44089e-16$ por cuestiones numéricas del algoritmo utilizado) y una cota superior igual a la anterior iteración.

ITERACIÓN n=1

CPLEX 20.1.0.0: optimal integer solution; objective $-4.440892099e-16$

0 MIP simplex iterations

0 branch-and-bound nodes

Cota inferior del regret óptimo alpha = $-4.44089e-16$

Solución actual $\hat{y}^n =$

y :=

```
1 2 0
1 4 1
2 3 0
2 5 0
3 6 0
4 2 0
4 5 0
4 7 1
5 3 0
5 6 0
5 8 0
6 9 0
7 5 0
7 8 1
8 6 0
8 9 1;
```

CPLEX 20.1.0.0: optimal integer solution; objective 1.2

0 MIP simplex iterations

0 branch-and-bound nodes

absmipgap = 4.44089e-16, relmipgap = 3.70074e-16

Máximo regret $R(\hat{x}^n, \hat{y}^n) = 1.200000$

Escenario más desfavorable $\hat{s}^n = (\hat{c}^n, \hat{f}^n)$

fbar :=

```
1 2 0.4
1 4 1
2 3 1
2 5 1
3 6 1
```

```

4 2  1
4 5  1
4 7  1
5 3  1
5 6  1
5 8  1
6 9  1
7 5  1
7 8  1
8 6  1
8 9  1.6;

```

Cota superior del regret óptimo $\beta = 1.2$

Para el problema de transporte entre origen y destino, en la segunda iteración obtenemos cómo solución la que pasa por el nodo 2, 5 y 8. En este caso los costes de los arcos con incertidumbre van a ser de 1.6 unidades ya que la solución pasa por dichos arcos y así obtenemos el peor escenario posible. Resolvemos el problema de separación y maestro obteniendo las correspondientes cotas del problema (*RND*), $\alpha = 0,6$ y $\beta = 1,2$.

ITERACIÓN $n=2$

CPLEX 20.1.0.0: optimal integer solution; objective 0.6

4 MIP simplex iterations

0 branch-and-bound nodes

Cota inferior del regret óptimo $\alpha = 0.6$

Solución actual $\hat{y}^n =$

```

y :=
1 2  1
1 4  0
2 3  0
2 5  1

```

```
3 6 0
4 2 0
4 5 0
4 7 0
5 3 0
5 6 0
5 8 1
6 9 0
7 5 0
7 8 0
8 6 0
8 9 1;
```

CPLEX 20.1.0.0: optimal integer solution; objective 1.2

0 MIP simplex iterations

0 branch-and-bound nodes

Máximo regret $R(\hat{x}^n, \hat{y}^n) = 1.200000$

Escenario más desfavorable $\hat{s}^n = (\hat{c}^n, \hat{f}^n)$

fbar :=

```
1 2 1.6
1 4 1
2 3 1
2 5 1
3 6 1
4 2 1
4 5 1
4 7 1
5 3 1
5 6 1
5 8 1
```

```
6 9 1
7 5 1
7 8 1
8 6 1
8 9 1.6;
```

Cota superior del regret óptimo $\beta = 1.2$

En la iteración tercera la solución propone pasar por los nodos 4, 5 y 6. Por lo tanto el peor escenario posible sería que los costes de los dos arcos con incertidumbre fuesen de 0.4 unidades. Obtenemos las mismas cotas α y β en la etapa anterior.

ITERACIÓN $n=3$

CPLEX 20.1.0.0: optimal integer solution; objective 0.6

1 MIP simplex iterations

0 branch-and-bound nodes

Cota inferior del regret óptimo $\alpha = 0.6$

Solución actual $y^{\hat{n}}$ =

$y :=$

```
1 2 0
1 4 1
2 3 0
2 5 0
3 6 0
4 2 0
4 5 1
4 7 0
5 3 0
5 6 1
5 8 0
6 9 1
```

```
7 5 0
7 8 0
8 6 0
8 9 0;
```

CPLEX 20.1.0.0: optimal integer solution; objective 1.2

0 MIP simplex iterations

0 branch-and-bound nodes

absmipgap = 4.44089e-16, relmipgap = 3.70074e-16

Máximo regret $R(\hat{x}^n, \hat{y}^n) = 1.200000$

Escenario más desfavorable $\hat{s}^n = (\hat{c}^n, \hat{f}^n)$

fbar :=

```
1 2 0.4
1 4 1
2 3 1
2 5 1
3 6 1
4 2 1
4 5 1
4 7 1
5 3 1
5 6 1
5 8 1
6 9 1
7 5 1
7 8 1
8 6 1
8 9 0.4;
```

Cota superior del regret óptimo $\beta = 1.2$

ITERACIÓN n=4

CPLEX 20.1.0.0: optimal integer solution; objective 1.2

8 MIP simplex iterations

0 branch-and-bound nodes

Cota inferior del regret óptimo alpha = 1.2

Solución óptima xsol :=

1 2 1 1

1 4 1 0

2 3 1 1

2 5 1 0

3 6 1 1

4 2 1 0

4 5 1 0

4 7 1 0

5 3 1 0

5 6 1 0

5 8 1 0

6 9 1 1

7 5 1 0

7 8 1 0

8 6 1 0

8 9 1 0;

ysol :=

1 2 1

1 4 0

2 3 1

2 5 0

3 6 1

4 2 0

4 5 0
 4 7 0
 5 3 0
 5 6 0
 5 8 0
 6 9 1
 7 5 0
 7 8 0
 8 6 0
 8 9 0;

La iteración cuatro es la última puesto que coinciden la cota inferior con la cota superior del valor objetivo óptimo del (RND), en este caso $\alpha = \beta = 1,2$. Esto era previsible puesto que en cuatro iteraciones se han incluido en el problema maestro los cuatro escenarios de coste posibles, por lo tanto, el último problema maestro resuelto equivale al problema original (RND) por incluir en las restricciones todos los escenarios posibles. Este hecho sucede por tratarse de un ejemplo con un número muy pequeño de escenarios lo que hace que no sea posible obtener cotas "finas" del objetivo óptimo antes de generar todos los costes posibles. A continuación damos el detalle del problema maestro resuelto en la última iteración.

La función objetivo a minimizar es r y vamos a ver a continuación las restricciones correspondientes. La primera de ellas corresponde al nodo de partida. Cómo partimos del 1, las únicas dos posibilidades para el transporte de unidades sería pasar al nodo 2 o 4. Por tanto lo reflejamos en la siguiente restricción :

subject to B1[1,1]:
 $x[1,2,1] + x[1,4,1] = 1;$

Por otro lado, al nodo 9 sólo se puede acceder desde el nodo 6 o 9 según la estructura de la red, es decir, la restricción de balance de flujo en el nodo 9 involucra a las variables $x_{6,9,1}$ y $x_{8,9,1}$ cómo se puede ver a continuación.

subject to B2[9,1]:

$$-x[8,9,1] - x[6,9,1] = -1;$$

Cuando nos encontramos en un nodo diferente al de origen o destino tenemos las siguientes restricciones dónde $b_{ip} = 0$ ya que ninguna unidad se queda en otro nodo distinto del nueve. La primera de ellas corresponde con todos los arcos que pueden llegar y salir del nodo dos. Por ejemplo podríamos llegar al nodo dos por el nodo uno o cuatro y pasaríamos al nodo tres o cinco de forma que si llegamos del nodo uno, no podríamos llegar del nodo cuatro y por tanto uno de los dos sumandos negativos debe ser cero. Análogamente a partir del nodo dos sólo podríamos ir al nodo tres o cinco y por tanto sólo uno de los dos sumandos positivos es cero y por tanto la suma final tiene que ser cero, y así sucesivamente para todos los nodos restantes.

$$\text{subject to B3[2,1]: } -x[1,2,1] + x[2,3,1] + x[2,5,1] - x[4,2,1] = 0;$$

$$\text{subject to B3[3,1]: } -x[2,3,1] + x[3,6,1] - x[5,3,1] = 0;$$

$$\text{subject to B3[4,1]: } x[4,5,1] - x[1,4,1] + x[4,7,1] + x[4,2,1] = 0;$$

$$\text{subject to B3[5,1]: } -x[4,5,1] + x[5,6,1] - x[2,5,1] + x[5,8,1] + x[5,3,1] - x[7,5,1] = 0;$$

$$\text{subject to B3[6,1]: } -x[5,6,1] - x[3,6,1] + x[6,9,1] - x[8,6,1] = 0;$$

$$\text{subject to B3[7,1]: } x[7,8,1] - x[4,7,1] + x[7,5,1] = 0;$$

$$\text{subject to B3[8,1]: } -x[7,8,1] + x[8,9,1] - x[5,8,1] + x[8,6,1] = 0;$$

Las siguientes restricciones corresponden con $x_{ijp} \leq y_{ij}$, equivalentemente $x_{ijp} - y_{ij} \leq 0$, es decir, si el diseño de la red no contiene el arco (i, j) no se envía flujo a través de dicho arco. Cómo sólo hay que conectar el nodo uno con el nueve, es decir el tercer índice de las variables x es siempre uno, $p = 1$, y (i, j) varía a lo largo de los arcos.

$$\text{subject to B4[1,2,1]: } x[1,2,1] - y[1,2] \leq 0;$$

$$\text{subject to B4[2,3,1]: } x[2,3,1] - y[2,3] \leq 0;$$

$$\text{subject to B4[4,5,1]: } x[4,5,1] - y[4,5] \leq 0;$$

$$\text{subject to B4[5,6,1]: } x[5,6,1] - y[5,6] \leq 0;$$

$$\text{subject to B4[7,8,1]: } x[7,8,1] - y[7,8] \leq 0;$$

$$\text{subject to B4[8,9,1]: } x[8,9,1] - y[8,9] \leq 0;$$

subject to B4[1,4,1]: $x[1,4,1] - y[1,4] \leq 0$;
 subject to B4[2,5,1]: $x[2,5,1] - y[2,5] \leq 0$;
 subject to B4[3,6,1]: $x[3,6,1] - y[3,6] \leq 0$;
 subject to B4[4,7,1]: $x[4,7,1] - y[4,7] \leq 0$;
 subject to B4[5,8,1]: $x[5,8,1] - y[5,8] \leq 0$;
 subject to B4[6,9,1]: $x[6,9,1] - y[6,9] \leq 0$;
 subject to B4[4,2,1]: $x[4,2,1] - y[4,2] \leq 0$;
 subject to B4[5,3,1]: $x[5,3,1] - y[5,3] \leq 0$;
 subject to B4[7,5,1]: $x[7,5,1] - y[7,5] \leq 0$;
 subject to B4[8,6,1]: $x[8,6,1] - y[8,6] \leq 0$;

Las siguientes restricciones se corresponden con los cortes de Benders $\sum_{(i,j) \in A} f_{ij}^q y_{ij} + \sum_{(i,j) \in A, p \in P} c_{ijp}^q x_{ijp} - D(\bar{x}^q, \bar{y}^q, s^q) \leq r, q = 0, \dots, n - 1$. Vamos a explicar el primero de ellos pues los demás son análogos. En este caso, los términos que acompañan a las x_{ijp} son los $c_{ijp} = 0,3$ para todos los arcos y los que acompañan a los y_{ij} son los $f_{ij} = 1$ en todos los arcos salvo en los dos con incertidumbre. En el peor escenario para la primera solución, el primer coste del arco del nodo uno al dos era 1.6 y el coste del arco del nodo ocho al nueve era de 0.4 y es por eso que en dichos arcos modificamos los f_{ij} . Además cuando vimos cuál era el escenario más desfavorable para la solución que pasaba por los nodos uno, dos, tres, seis y nueve vimos que el coste de construcción de todos los arcos salvo el primero sería de una unidad puesto que no hay incertidumbre mientras que el restante sería de 1.6, que eso daría un coste total de 4.6 unidades y por tanto tenemos el coste determinístico $D(\bar{x}^1, \bar{y}^1, s^1)$ que aparece en el lado derecho de la primera restricción.

subject to CUTS[1]:

$0.3*x[1,2,1] + 0.3*x[2,3,1] + 0.3*x[4,5,1] + 0.3*x[5,6,1] +$
 $0.3*x[7,8,1] + 0.3*x[8,9,1] + 0.3*x[1,4,1] + 0.3*x[2,5,1] +$
 $0.3*x[3,6,1] + 0.3*x[4,7,1] + 0.3*x[5,8,1] + 0.3*x[6,9,1] +$
 $0.3*x[4,2,1] + 0.3*x[5,3,1] + 0.3*x[7,5,1] + 0.3*x[8,6,1] + 1.6*y[1,2]$
 $+y[2,3] +y[4,5] +y[5,6] + y[7,8] + 0.4*y[8,9] + y[1,4] + y[2,5] +$

$y[3,6] + y[4,7] + y[5,8] + y[6,9] + y[4,2] + y[5,3] + y[7,5] + y[8,6] - r \leq 4.6;$

subject to CUTS[2]:

$0.3*x[1,2,1] + 0.3*x[2,3,1] + 0.3*x[4,5,1] + 0.3*x[5,6,1] +$
 $0.3*x[7,8,1] + 0.3*x[8,9,1] + 0.3*x[1,4,1] + 0.3*x[2,5,1] +$
 $0.3*x[3,6,1] + 0.3*x[4,7,1] + 0.3*x[5,8,1] + 0.3*x[6,9,1] +$
 $0.3*x[4,2,1] + 0.3*x[5,3,1] + 0.3*x[7,5,1] + 0.3*x[8,6,1] + 0.4*y[1,2]$
 $+ y[2,3] + y[4,5] + y[5,6] + y[7,8] + 1.6*y[8,9] + y[1,4] + y[2,5] +$
 $y[3,6] + y[4,7] + y[5,8] + y[6,9] + y[4,2] + y[5,3] + y[7,5] + y[8,6]$
 $- r \leq 4.6;$

subject to CUTS[3]:

$0.3*x[1,2,1] + 0.3*x[2,3,1] + 0.3*x[4,5,1] + 0.3*x[5,6,1] +$
 $0.3*x[7,8,1] + 0.3*x[8,9,1] + 0.3*x[1,4,1] + 0.3*x[2,5,1] +$
 $0.3*x[3,6,1] + 0.3*x[4,7,1] + 0.3*x[5,8,1] + 0.3*x[6,9,1] +$
 $0.3*x[4,2,1] + 0.3*x[5,3,1] + 0.3*x[7,5,1] + 0.3*x[8,6,1] + 1.6*y[1,2]$
 $+ y[2,3] + y[4,5] + y[5,6] + y[7,8] + 1.6*y[8,9] + y[1,4] + y[2,5] +$
 $y[3,6] + y[4,7] + y[5,8] + y[6,9] + y[4,2] + y[5,3] + y[7,5] + y[8,6]$
 $- r \leq 5.2;$

subject to CUTS[4]:

$0.3*x[1,2,1] + 0.3*x[2,3,1] + 0.3*x[4,5,1] + 0.3*x[5,6,1] +$
 $0.3*x[7,8,1] + 0.3*x[8,9,1] + 0.3*x[1,4,1] + 0.3*x[2,5,1] +$
 $0.3*x[3,6,1] + 0.3*x[4,7,1] + 0.3*x[5,8,1] + 0.3*x[6,9,1] +$
 $0.3*x[4,2,1] + 0.3*x[5,3,1] + 0.3*x[7,5,1] + 0.3*x[8,6,1] + 0.4*y[1,2]$
 $+ y[2,3] + y[4,5] + y[5,6] + y[7,8] + 0.4*y[8,9] + y[1,4] + y[2,5] +$
 $y[3,6] + y[4,7] + y[5,8] + y[6,9] + y[4,2] + y[5,3] + y[7,5] + y[8,6]$
 $- r \leq 4;$

Cómo se puede ver en este ejemplo, el pequeño número de escenarios considerados hace que el algoritmo 1 los introduzca todos como cortes en el problema maestro.

3. Estudio experimental.

El objetivo del estudio es comparar un diseño de red *minmax regret* con el diseño óptimo bajo costes esperados. Hay que tener en cuenta que para obtener esta segunda red es necesario conocer la distribución aleatoria de los costes o, al menos, su media. Obviamente, si la eficiencia de cada una de estas redes es cuantificada durante un largo período de tiempo en el que las distribuciones estadísticas se mantienen inalterables, la Ley Fuerte de los Grandes Números otorgará el mejor comportamiento promedio al diseño bajo costes esperados. Sin embargo, en la práctica estas condiciones distribucionales suelen ser dinámicas (costes de mantenimiento creciente, fluctuación de demanda,...) o bien, se ven afectados por eventos incontrolables.

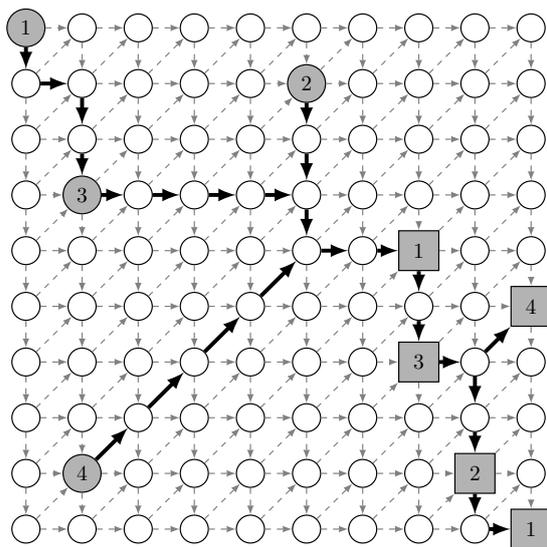


Figura 9: Solución óptima bajo costes esperados.

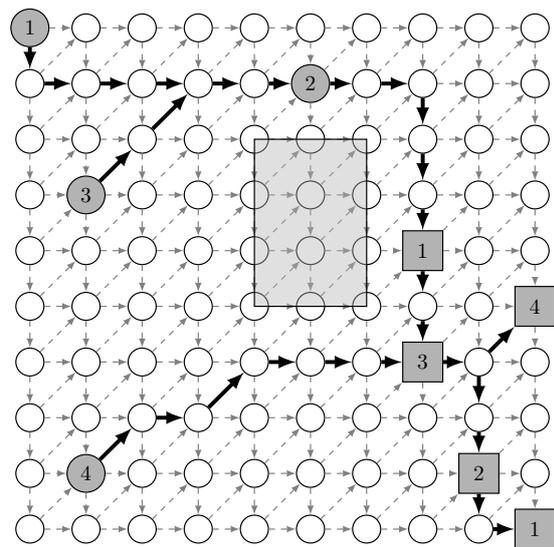


Figura 10: Solución *minmax regret* tomando cómo intervalo de incertidumbre el soporte de los costes

Supondremos que los costes de mantenimiento por arco son de una unidad y los costes de transporte son 0.3 unidades fuera de la zona de incertidumbre. Dentro del rectángulo, los costes de mantenimiento de cada arco se encuentran en el intervalo $[0,4,1,6]$ que se tomará como soporte para la distribución de costes aleatorios con los que se compara.

El experimento consiste en comparar el diseño *minmax regret* frente al diseño bajo

costes esperados. Se evalúa el coste de cada uno de los diseños un número n de veces. Para ello, se obtienen n realizaciones de los costes de mantenimiento individuales en los arcos y se calcula el coste total de la red. El coste del diseño *minmax regret* es constante porque no se verá afectado por la variación de costes de las diferentes realizaciones, por tanto obtenemos una muestra constante, mientras que la otra es una muestra aleatoria simple dónde podremos usar las herramientas de inferencia.

Para poder realizar dicha comparación realizaremos un contraste de hipótesis de igualdad de medias, es decir, $H_0 : \mu = \mu_0$ frente a $H_1 : \mu \neq \mu_0$. Cómo los costes de construcción de los arcos con incertidumbre siguen distribuciones generales como puede ser suma de betas dependientes, suma de betas con uniformes, . . . , usaremos el siguiente intervalo de confianza asintótico para distribuciones desconocidas a nivel $1 - \alpha$:

$$\left(\bar{X} - \frac{S}{\sqrt{n}} Z_{1-\frac{\alpha}{2}}, \bar{X} + \frac{S}{\sqrt{n}} Z_{1-\frac{\alpha}{2}} \right)$$

dónde $Z_{1-\frac{\alpha}{2}}$ representa el cuantil $1 - \frac{\alpha}{2}$ de una $N(0, 1)$, S^2 es la varianza muestral y n el tamaño muestral. Este es un intervalo de confianza asintótico dónde hay que tener en cuenta que para n suficientemente grande, la primera solución tendrá un coste total inferior al de la segunda por la *Ley Fuerte de los Grandes Números*. Con este experimento se pretende poner de manifiesto dos ideas importantes:

1. Incluso si existen datos que permitan estimar el comportamiento distribucional de los costes ambas soluciones pueden ser comparables.
2. La solución *minmax regret* es adecuada si su frecuencia de uso no es muy alta en el tiempo. (LFGN)

Debido a la incertidumbre sobre los costes de los arcos dentro de la zona rectangular, puede que la única forma de abordar el problema sea con la técnica *minmax regret*, en particular si desconocemos la distribución de los costes de los arcos. Sin embargo a medida que vamos aumentando el grado de conocimiento del comportamiento distribucional para los costes, la función objetivo del diseño bajo costes esperados tiende a ser

menor que el de la solución *minmax regret*. Debido principalmente a la *Ley Fuerte de los Grandes Números* ya que a medida que el tamaño muestral va incrementando juega a favor de la primera solución. La muestra aleatoria va generando valores de los costes de los arcos cercanos a la media, que en este caso es uno, y si hay algunos costes bastante alejados de la media se contrarrestan unos con otros dando como resultado un coste total promedio que va a ser menor que el coste total de la segunda figura. Sin embargo, en este estudio experimental vamos a tener un tamaño muestral pequeño para poder comparar las dos soluciones.

Con la ayuda de *RStudio*, supondremos que los costes de todos los arcos en la zona rectangular tienen incertidumbre, pero en el experimento sólo cinco de dichos arcos pertenecen a la solución (véase figura 9) y los denotaremos a, b, c, d, e . En el experimento supondremos que los costes de los arcos con incertidumbre siguen distintas distribuciones pero siempre manteniendo la media, por tanto el diseño de red bajo costes esperados no cambiará. Además dentro de cada modelo distribucional se analizarán distintos tamaños muestrales para comprobar la estabilidad del diseño *minmax regret*.

Realizaremos en todas las secciones siguientes el contraste $H_0 : \mu = 44,4$ frente a $H_1 : \mu \neq 44,4$, ya que 44.4 es el coste total de la solución *minmax regret*. Rechazaremos la hipótesis nula si $\mu_0 = 44,4$ no pertenece al intervalo de confianza asintótico anterior. Recordemos que *RStudio* tiene predeterminado calcular la cuasivarianza muestral con el comando $var(x)$ y es por ello que multiplicamos por $\frac{N-1}{N}$ para obtener la varianza muestral, cómo se indica en la figura 11.

Además para tener más evidencia de rechazo o aceptación de H_0 calculamos el p-valor que es el mínimo entre α_1 y α_2 dónde:

$$\alpha_1 = 2\phi\left(\sqrt{n}\frac{\bar{X}_n - \mu_0}{S}\right) \quad \alpha_2 = 2\left(1 - \phi\left(\sqrt{n}\frac{\bar{X}_n - \mu_0}{S}\right)\right)$$

con ϕ la función de distribución de una $N(0, 1)$. Para el intervalo de confianza usaremos un nivel de significación del 5%.

3.1. Experimento con costes independientes

3.1.1. Distribución uniforme

En este caso supondremos las siguientes hipótesis distribucionales:

1. Todos los costes de los arcos en la zona de incertidumbre van a seguir una distribución uniforme en el intervalo $[0,4, 1,6]$, es decir, $X \sim U(0,4, 1,6)$.
2. Los costes de los arcos se distribuyen de forma independiente.

Consideramos una muestra aleatoria independiente de tamaño $N = 50$ para calcular el intervalo de confianza asintótico con nivel de significación del 5%. En la figura 11, los valores a,b,c,d,e de las líneas 15-19 corresponden a los costes de los cinco arcos del diseño de red bajo costes esperados que pertenecen al rectángulo de incertidumbre, (véase figura 9).

```
1 #Distribución uniforme independiente
2
3 set.seed(1234)
4
5 #tamaño muestral:
6
7 N =50;
8
9 #Inicialización del vector de costes:
10
11 coste = rep(0,N)
12
13 for (n in 1:N) {
14
15 a = runif(1, min = 0.4, max =1.6)
16 b = runif(1, min = 0.4, max =1.6)
17 c = runif(1, min = 0.4, max =1.6)
18 d = runif(1, min = 0.4, max =1.6)
19 e = runif(1, min = 0.4, max =1.6)
20
21 coste[n] = a + b +c + d + e + 20 +(0.3*58)}
22
23 xmedia = sum(coste)/N;
24 alpha = 0.05
25 Z =qnorm(1-alpha/2, mean = 0, sd =1, lower.tail = TRUE, log.p = FALSE)
26 cuasivar = var(coste)
27 varmuestral = (N-1)/N*cuasivar
28 cuasidesv = sqrt(varmuestral)
29 x = xmedia - cuasidesv/sqrt(N)* Z
30 y = xmedia + cuasidesv/sqrt(N)* Z
31
32 #Intervalo de confianza con nivel de significación 5% : (42.11946 ,42.59901)
33
```

Figura 11: Implementación del intervalo de confianza asintótico con costes según una distribución uniforme.

El intervalo de confianza es $(42,11946, 42,59901)$ que no incluye el valor 44.4 por lo que rechazaríamos la hipótesis nula, es decir, la igualdad de medias y además hemos calculado el p-valor correspondiente que nos da más evidencias de rechazo ya que es prácticamente cero.

Esto es debido a que los valores de la muestra de los costes con incertidumbre suelen compensarse bajo la condición de independencia. Por tanto, bajo este modelo distribucional los costes del diseño de la figura 9 son bastantes estables cómo demuestra la reducida amplitud del intervalo de confianza.

Se representan los siguientes histogramas de la muestra de costes en la figura 12 variando el tamaño muestral N . El eje x representará los distintos valores de los costes totales de la red de la figura 9, mientras que el eje y representa la frecuencia absoluta con la que obtenemos dicho valor.

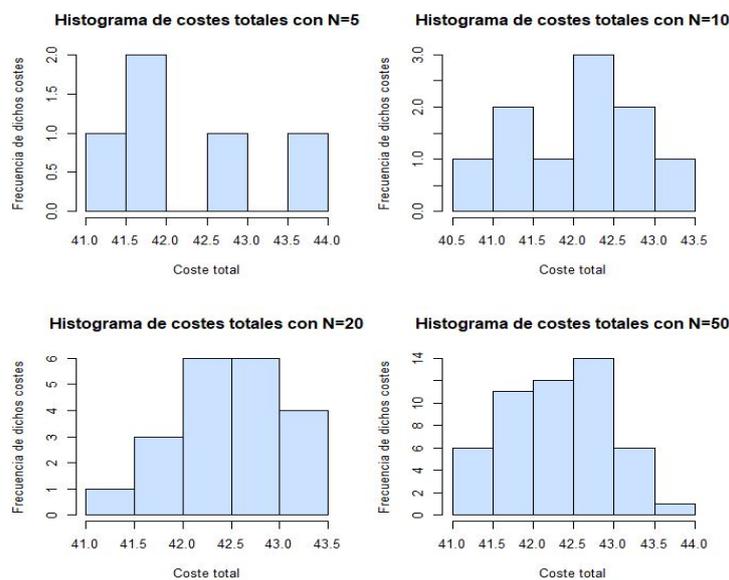


Figura 12: Histogramas representando el coste total de la solución según diferentes muestras.

Cómo se puede ver, a medida que aumenta el tamaño muestral, los costes se van concentrando en torno al valor 42.4 que corresponde al coste esperado del diseño de red de la figura 9. En cualquier caso el valor 44.4 se encuentra bastante alejado de los costes totales del problema de transporte por lo que rechazaríamos la igualdad de medias.

3.1.2. Distribución Beta

En este caso tenemos las siguientes hipótesis distribucionales:

1. Los costes de los arcos dentro de la zona de incertidumbre siguen una distribución

Beta con parámetros α y β , $X \sim B(\alpha, \beta)$.

2. Los costes de los arcos se distribuyen de forma independiente.
3. Mantener la media igual a uno y el soporte del coste en $[0,4, 1,6]$. Cómo sabemos que la esperanza de una distribución beta es $\frac{\alpha}{\alpha+\beta}$ entonces:

$$E(0,4 + 1,2X) = 1$$

$$0,4 + 1,2 E(X) = 1$$

$$0,4 + 1,2 \frac{\alpha}{\alpha + \beta} = 1$$

Veamos en la gráfica de la figura 13 que para valores suficientemente pequeños de los parámetros se va a favorecer la solución robusta.

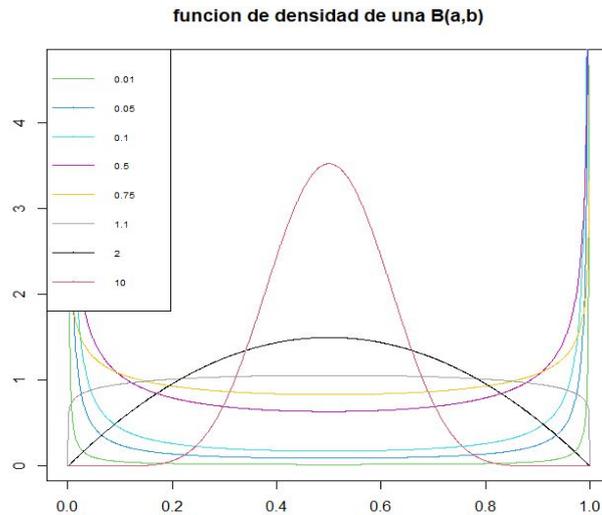


Figura 13: Función de densidad de betas suponiendo $\alpha = \beta$.

Esto es debido a que el área se concentra en los extremos para α y β menores que uno, mientras que cuando dichos parámetros son mayores, la función de densidad concentra la masa de probabilidad en el centro del intervalo. Por tanto acabará favoreciendo a la solución bajo costes promedios. Modificando los parámetros α y β obtenemos los intervalos de confianza de la tabla 1:

	alpha	beta	Intervalo de	confianza	p-valor
Experimento 1	0.00010	0.00010	40.99674	43.08126	8.902420e-06
Experimento 2	0.00050	0.00050	41.71429	43.08571	1.086602e-08
Experimento 3	0.00075	0.00075	42.03526	43.72474	4.207865e-04
Experimento 4	0.00100	0.00100	43.46219	44.69781	2.495390e-01
Experimento 5	0.01000	0.01000	41.56851	43.64119	6.361449e-03
Experimento 6	0.05000	0.05000	41.84120	43.37917	7.509668e-03
Experimento 7	0.10000	0.10000	41.77726	43.14163	6.700911e-11
Experimento 8	0.50000	0.50000	41.99779	42.84573	7.744837e-09
Experimento 9	0.75000	0.75000	42.20500	43.00117	8.337897e-07

Cuadro 1: Tabla de los intervalos de confianza y p-valores de los costes según los parámetros α y β con $N = 10$.

Cómo vemos el intervalo de confianza muy pocas veces contiene el valor 44.4 y además tenemos p-valores en la mayoría de casos excesivamente pequeños, por lo que rechazaríamos la igualdad de medias.

3.2. Experimento con costes relacionados.

3.2.1. Distribución uniforme discreta.

En la mayoría de las situaciones prácticas, los costes involucrados en el transporte y mantenimiento de la red interactúan entre ellos, es decir, no funcionan de forma independiente cómo se ha modelado en los experimentos anteriores. Es evidente que un aumento en el coste del combustible afecta a todos los costes de transporte pero también afecta a los costes de mantenimiento de las diferentes conexiones de la red. En lo que sigue incluimos este importante factor en el experimento de comparación del diseño *minmax regret* respecto al diseño bajo costes esperados y veremos que afecta de forma sensible a los resultados obtenidos previamente.

El primer modelo distribucional que analizaremos tiene las siguientes hipótesis:

1. Los arcos con incertidumbre siguen una distribución uniforme discreta, $X \sim UD(0, 1)$.

En el caso $X = 0$ los costes de los arcos con incertidumbre van a tener un coste

de mantenimiento de 0.4 unidades y si $X = 1$ de 1.6 y toman dichos valores con probabilidades $\{0,5, 0,5\}$.

2. Los arcos que intervienen en la solución van a tener todos el mismo coste de mantenimiento de acuerdo al valor generado en el punto anterior.

Mostramos los intervalos de confianza correspondientes con un nivel de significación del 5% como sigue:

	Intervalo de	confianza	p-valor
N=5	42.09635	46.30365	0.852178900
N=10	42.71249	45.68751	0.792147400
N=15	41.51249	44.48751	0.065086730
N=17	41.15286	44.00008	0.012054370
N=20	41.39181	44.00819	0.010866050
N=25	41.59252	43.92748	0.005901070
N=30	41.94817	44.05183	0.009087467
N=50	42.18526	43.81474	0.000757495

Cuadro 2: Tabla recogiendo los intervalos de confianza y p-valores de los costes.

Notemos que hasta tamaño muestral $N = 20$ podemos aceptar la igualdad de medias con error de tipo I de 0.01 pues el valor 44.4 se encuentra dentro del intervalo. A medida que vamos incrementando N , μ_0 deja de estar dentro del intervalo. El caso más claro de rechazo es para $N = 50$ pues dicho p-valor es bastante cercano a 0 y el intervalo de confianza se encuentra alejado del valor 44.4.

Ahora vamos a representar mediante histogramas el comportamiento de los costes totales bajo la hipótesis distribucional $UD(0, 1)$ con dependencia total entre costes para $N = 50$. Notemos que el coste total sólo puede tener dos valores, dependiendo de si $X = 0$ o $X = 1$ con probabilidades $\frac{1}{2}$. Sin embargo, para esta muestra en concreto la cantidad de veces que el coste del arco toma el valor 0.4 es bastante superior al de 1.6.

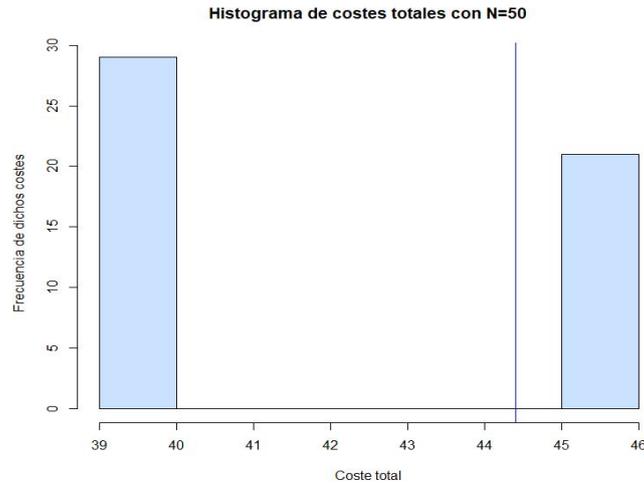


Figura 14: Histograma de los costes totales según una distribución uniforme discreta.

El resultado obtenido es completamente diferente comparado con el caso de costes independientes ya que en este caso observamos que si las condiciones distribucionales no son estables entonces la solución *minmax regret* es competitiva frente a la solución óptima bajo costes esperados.

3.2.2. Dependencia de los arcos con distribución uniforme discreta según su orientación.

En este nuevo experimento numérico, suponemos las siguientes hipótesis distribucionales:

1. Los costes de los arcos con incertidumbre siguen una distribución uniforme discreta, $X \sim UD(0,4, 1,6)$ con $P(X = 0,4) = \frac{1}{2}$ y $P(X = 1,6) = \frac{1}{2}$.
2. Mantenemos la media de los costes de los arcos en una unidad.
3. Los arcos oblicuos y verticales van a tener el mismo coste generado desde la misma muestra aleatoria simple, mientras que todos los costes de los arcos horizontales tienen el mismo coste generado por otra muestra aleatoria simple diferente.

La línea roja representada en la figura 15 es el umbral a partir del cual aceptamos H_0 .

3.2.3. Dependencia de los arcos según una distribución uniforme.

En este experimento supondremos:

1. Los costes de los arcos con incertidumbre siguen una uniforme en el intervalo $(0,4, 1,6)$, es decir, $X \sim U(0,4, 1,6)$.
2. Todos los arcos de la solución tienen el mismo coste según la misma muestra aleatoria simple.
3. Mantenemos la media igual a uno.

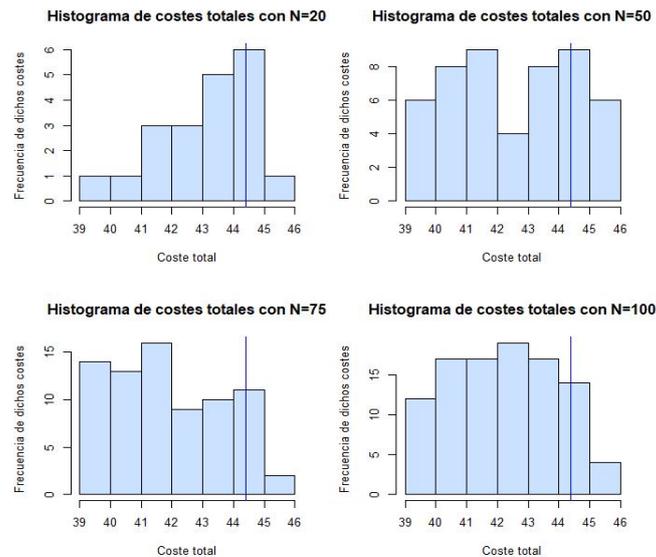


Figura 16: Histograma de los costes totales junto a su frecuencia absoluta variando el tamaño muestral N .

Cómo vemos el valor 44.4 está en un intervalo de clase que tiene una frecuencia bastante alta incluso aumentando el tamaño muestral, a diferencia del caso de los costes con distribución uniforme según una muestra independiente. Aún así la media de los costes no parece que sea de 44.4 pues queda bastante desplazado a la derecha, para poder confirmar el rechazo de la igualdad de medias para tamaños muestrales grandes haremos

la nube de puntos de los correspondientes p-valores que ilustramos a continuación con una línea roja que marcará el umbral a partir del cual podemos aceptar la hipótesis nula con un nivel de significación del 5%.

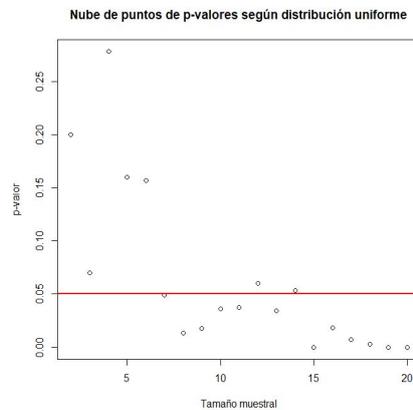


Figura 17: Nube de puntos de los p-valores.

3.2.4. Dependencia de los arcos con distribución uniforme según su orientación.

Tenemos las siguientes hipótesis distribucionales:

1. Los arcos dentro de la zona rectangular con incertidumbre suponemos que siguen una distribución uniforme en el intervalo $(0,4, 1,6)$, es decir, $X \sim U(0,4, 1,6)$
2. La dependencia de dichos costes de arcos va a venir dada por su posición. Los arcos que se encuentren horizontalmente tendrán un cierto coste mientras que los de posición vertical y oblicua tendrán el mismo coste generado por la misma muestra aleatoria.
3. Mantenemos la media igual a uno.

Recogemos los correspondientes intervalos de confianza junto a los p-valores incrementando el tamaño muestral N y con un nivel de significación del 5%.

	Intervalo de	confianza	p-valor
N=2	40.91917	44.35333	4.408937e-02
N=3	41.73180	44.41452	5.253273e-02
N=4	40.63760	43.95486	1.291858e-02
N=5	41.63229	44.56882	8.260936e-02
N=6	42.05148	44.66926	1.195254e-01
N=7	42.01522	44.60284	9.839576e-02
N=8	42.04239	44.34554	4.010552e-02
N=9	41.89971	44.08092	1.129617e-02
N=10	42.05037	44.02064	6.633107e-03
N=15	42.21253	43.60048	2.464436e-05
N=20	42.22354	43.32757	8.034818e-09

Cuadro 4: Tabla con intervalos de confianza y p-valores.

A continuación modificamos una de las hipótesis distribucionales anteriores: los costes de los arcos horizontales y verticales son del mismo valor generado por la misma muestra aleatoria según una distribución uniforme en el intervalo $(0,4,1,6)$, es decir, $X \sim U(0,4,1,6)$, mientras que los arcos oblicuos tendrán un coste diferente generados por otra muestra. En la figura 18 representamos la muestra de costes totales de la red de la figura 9 mediante un histograma.

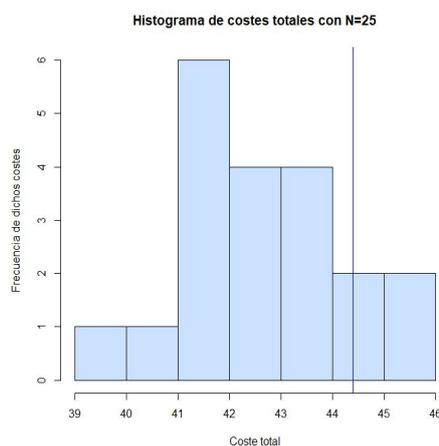


Figura 18: Histograma de coste totales con $N = 25$.

3.2.5. Costes dependientes con distribución beta.

Las hipótesis distribucionales son:

1. Los costes de mantenimiento de los arcos dentro de la zona de incertidumbre siguen una distribución beta $B(\alpha, \beta)$ trasladada al soporte $[0,4, 1,6]$ y manteniendo el coste medio en uno.
2. Los costes de los arcos van a ser generados por la misma muestra aleatoria de la distribución beta.

Variamos los parámetros α y β para calcular los correspondientes intervalos de confianza recogidos en la siguiente tabla:

	alpha	beta	Intervalo de	confianza
Experimento 1	0.0001	0.0001	42.09635	46.30365
Experimento 2	0.0010	0.0010	42.09635	46.30365
Experimento 3	0.0050	0.0050	42.09635	46.30365
Experimento 4	0.0075	0.0075	42.09635	46.30365
Experimento 5	0.0100	0.0100	40.42356	45.57644
Experimento 6	0.0500	0.0500	40.42363	45.57644
Experimento 7	0.1000	0.1000	40.22043	44.95385
Experimento 8	0.2500	0.2500	42.20866	45.80195
Experimento 9	0.5000	0.5000	39.55585	43.25322
Experimento 10	1.0000	1.0000	41.18519	43.73639

Cuadro 5: Intervalos de confianza según los parámetros α y β con $N = 5$.

Cómo hemos mencionado anteriormente, los parámetros α y β que harán aceptar H_0 con mayor frecuencia son los más pequeños que hacen que los costes generados se concentren en torno a los extremos del intervalo soporte $[0,4, 1,6]$. Por tanto dichos parámetros favorecerán la solución robusta y tendremos que aceptar la igualdad de medias. Según los intervalos de confianza de la tabla anterior mostramos que para valores de los parámetros superiores a uno no podremos aceptar la hipótesis nula. Además hemos calculado los p-valores ilustrados en la figura 19 de nube de puntos:

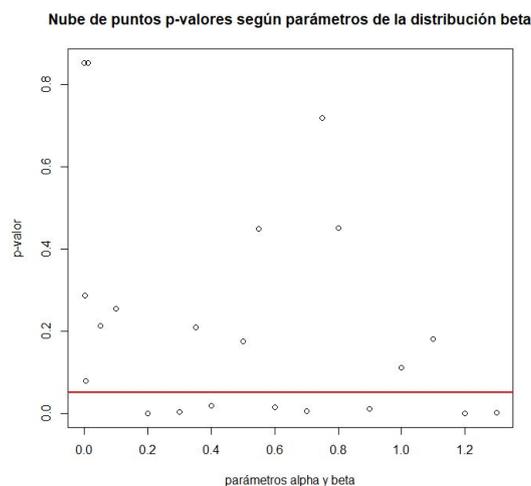


Figura 19: Nube de puntos de los p-valores para $N = 5$.

En esta gráfica hemos representado los p-valores correspondientes de variar los parámetros α y β . Notemos que a partir del parámetro con valor uno el p-valor se sitúa por debajo de la línea roja que representa el nivel de significación del 5%, es decir, los puntos que se encuentren por debajo de dicha línea serán menores que 0.05 y por tanto tendríamos que rechazar la igualdad de medias. En cambio para valores muy pequeños de dichos parámetros tenemos p-valores bastante altos y por tanto aceptaríamos la hipótesis nula.

3.2.6. Dependencia del coste según la orientación del arco.

Suponemos las siguientes hipótesis distribucionales:

1. Los costes de los arcos con incertidumbre siguen una distribución beta, $B(\alpha, \beta)$ trasladada al soporte $[0, 4, 1, 6]$.
2. Los costes de arcos con posición horizontal van a ser los mismos mientras que los costes verticales y oblicuos serán iguales, todos ellos siguiendo una distribución beta a la que iremos modificando sus parámetros α y β .
3. Mantener la media igual a uno.

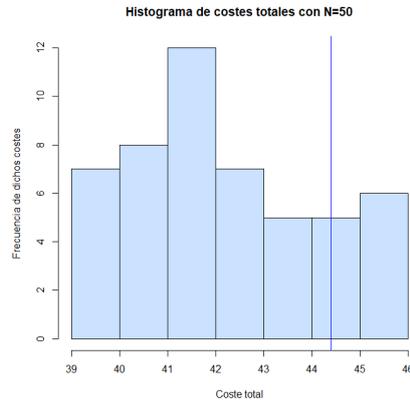


Figura 20: Histograma costes totales con $N = 50$.

En la figura 20 hemos considerado α y β suficientemente pequeños como para que la masa de probabilidad de la función de densidad se concentre en los extremos. Hemos señalado con una línea azul el valor 44.4 para poder ver en cuántos casos de $N = 50$ aceptaríamos la igualdad de medias. Además hemos calculado los p-valores recogidos en la tabla 6:

	alpha	beta	p-valor
Experimento 1	0.0001	0.0001	0.2868687000
Experimento 2	0.0010	0.0010	0.7680618000
Experimento 3	0.0050	0.0050	0.0274281500
Experimento 4	0.0075	0.0075	0.1150181000
Experimento 5	0.0100	0.0100	0.4026598000
Experimento 6	0.0500	0.0500	0.1933277000
Experimento 7	0.1000	0.1000	0.1420983000
Experimento 8	0.2500	0.2500	0.1723845000
Experimento 9	0.5000	0.5000	0.0008663385

Cuadro 6: P-valores según los parámetros α y β con $N = 5$.

Para aquellos p-valores menores que 0.05 rechazamos la hipótesis nula con un nivel de significación del 5%. Por tanto para los parámetros mayores que 0.5 en este caso tendremos que rechazar la igualdad de medias.

3.3. Conclusiones.

En este estudio experimental hemos considerado dos casos diferentes, experimentos con costes independientes y dependientes. En el primero de ellos obtenemos que en la mayoría de experimentos rechazamos la igualdad de medias de los costes imputados a las dos redes comparadas 3.1.1, 3.1.2. Esto es debido a que los costes suelen compensarse, por tanto la solución robusta no suele ser competitiva. Además si repitiéramos un número grande de veces dicho experimento, se apreciaría mejor la sensibilidad del criterio *minmax regret*, que en este caso es deficiente ya que por la Ley Fuerte de los Grandes Números sabemos que se favorece la solución bajo costes promedios.

En el caso dónde los costes son dependientes, con mayor frecuencia aceptamos la igualdad de medias: 3.2.1,3.2.2, 3.2.3, 3.2.4, 3.2.5,3.2.6. Esto es debido a que todos los costes de los arcos con incertidumbre toman el mismo valor según la muestra o en otros casos tienen el mismo coste según la orientación de dichos arcos. Por tanto la solución robusta es competitiva pues si todos los costes son iguales y toman un valor próximo a 1.6 entonces el coste total de la solución será cercano a 44.4 y por tanto podremos aceptar la igualdad de medias.

En definitiva, podemos decir que cuando las condiciones distribucionales no son estables y existen dependencias entre los costes involucrados la solución *minmax regret* es competitiva frente a la solución óptima bajo costes esperados.

Referencias

- [1] C. Albarrán (2020), “*Las empresas imponen nuevos requisitos de conectividad*”, Redes&Telecom.

<https://www.redestelecom.es/conectividad/noticias/1121577051003/empresas-imponen-nuevos-requisitos-de-conectividad.1.html>

- [2] C.G. Bolinches (2021), *El caos en el transporte marítimo que provoca problemas en el comercio mundial*, elDiario.es. https://www.eldiario.es/economia/caos-transporte-maritimo-provoca-problemas-comercio-mundial_1_8422517.html

- [3] E.C. Brands (2021), *Por qué son necesarias las interconexiones (y sin embargo España está a la cola de Europa)*, Cotizalia.

https://www.elconfidencial.com/empresas/2021-11-30/interconexiones-internacionales-europa-bra_3299944/

- [4] Cluster Industrial (2020), *Los beneficios de una infraestructura robusta en el uso de sensores en la industria*. <https://www.clusterindustrial.com.mx/noticia/4142/los-beneficios-de-una-infraestructura-robusta-en-el-uso-de-sensores-en-la-industria-4>

- [5] E. Conde, M. Leal (Diciembre 2021), *A robust optimization model for distribution network design under a mixed integer set of scenarios*, *Computers and Operations Research*, Volume 136, 105493.

- [6] O. Kardoudi (2022), *La red que puede ahorrar el 48 % de la factura eléctrica*, El Confidencial. https://www.elconfidencial.com/tecnologia/novaceno/2022-02-14/la-red-europea-que-puede-ahorrar-un-48-del-coste-de-la-electricidad_3375602/

- [7] H. Montero (2022), *El gas para 70 años que España no quiere explotar*, La Razón.

<https://www.larazon.es/economia/20220312/6ul3guybujggfcwnffht4qulu.html>