



UNIVERSIDAD DE CHILE
FACULTAD DE CIENCIAS FÍSICAS Y MATEMÁTICAS
DEPARTAMENTO DE INGENIERÍA MATEMÁTICA

ESTIMACIÓN DÉBIL DE LA SENSIBILIDAD DEL OBJETIVO EN PROBLEMAS LINEALES

TESIS PARA OPTAR AL TÍTULO DE MAGISTER
EN GESTIÓN DE OPERACIONES

MEMORIA PARA OPTAR AL TÍTULO DE INGENIERO
CIVIL MATEMÁTICO

RODRIGO ANDRÉS LÓPEZ INSINILLA

PROFESOR GUÍA:
FERNANDO ORDOÑEZ PIZARRO

MIEMBROS DE LA COMISIÓN:
ROBERTO COMINETTI COTTI-COMETTI
DANIEL ESPINOZA GONZALEZ
JORGE VERA ANDREO

SANTIAGO DE CHILE
DICIEMBRE 2011

ESTIMACIÓN DEBIL DE LA SENSIBILIDAD DEL OBJETIVO
EN PROBLEMAS LINEALES

RODRIGO ANDRÉS LÓPEZ INSINILLA

21 DE DICIEMBRE DE 2011

*a tu luz, que me despertó el 2010.
Gracias.*

RESUMEN DE LA TESIS
 PARA OPTAR AL TÍTULO DE
 INGENIERO CIVIL MATEMÁTICO
 POR: RODRIGO LÓPEZ INSINILLA
 FECHA: XX DE ENERO DE 2012
 PROFESOR GUÍA:
 Sr. FERNANDO ORDOÑEZ

En general, un problema de decisión $\min_{x \in F^0} f(x)$ está sometido a una gran cantidad de factores que pueden provocar incertidumbre respecto a la fidelidad de los valores de los datos que definen F^0 , causando que la respuesta de este no sea del todo confiable. Existen diversos métodos para hacerse cargo de la incerteza en los datos, como el Análisis de Escenarios, la Optimización Estocástica y la Simulación, entre otras. Si el decidor es adverso al riesgo, por ejemplo en situaciones donde las decisiones son poco frecuentes o bien las consecuencias de una mala decisión ponen en riesgo la vida de personas, la Optimización Robusta, es la estrategia que le permite ser en extremo conservador, buscando soluciones óptimas que sean factibles bajo cualquier escenario posible de datos. Lamentablemente un algoritmo robusto puede consumir vastos recursos computacionales. Resulta interesante ser capaz de predecir cuánto se *arriesga* (en términos de la función objetivo), al utilizar una solución *económica* que ignora la incertidumbre en vez de una *costosa* solución robusta, o dicho de otra forma, cuánto *cuesta* una solución conservadora en relación al problema con datos estimados (fácil de resolver). Es posible acotar este valor Δ , en términos de la sensibilidad estructural del problema, una característica intrínseca de la modelación, y el nivel de incertiza que al que estan sometidos los datos, de la siguiente forma:

$$\Delta \leq \underbrace{\left(\frac{2}{\kappa_{\mathcal{U}} + 1} \right)}_{\text{sensibilidad del problema}} \underbrace{\left(\max_{x \in F^0} f(x) - \min_{x \in F^0} f(x) \right)}_{\text{tamaño del problema}}$$

Donde $\kappa_{\mathcal{U}}$ es una medida llamada *Margen de Factibilidad* propuesta por Ben-Tal y Nemirovski, en situaciones donde la variabilidad de los datos puede ser modelada a través de un conjunto de incerteza \mathcal{U} poliedral. Ellos presentan una cota superior para Δ y en este trabajo se construye un modelo linearizado para computar una estimación simplificada de esta cota para problemas lineales con incertidumbre en la matriz de restricciones de desigualdad, descrita a través de un conjunto poliedral. Se aplicó este modelo a 16 problemas de la librería **NETLib**, asumiendo perturbaciones independientes de los parámetros considerados como *inciertos*. La estimación implementada consiguió buenas cotas ajustadas: Para un nivel de incerteza del 1% las cotas fueron, salvo por dos ocasiones, a lo más 6 veces el valor a estimar y en general el error de la estimación no supero el 8% del valor óptimo nominal. En estos problemas se pudo observar que el error en la cota estimada es proporcional al nivel de incerteza, de comprobarse esta idea, se presentaría una ventaja significativa al momento de estudiar el impacto sobre problemas con nivel de incerteza desconocido.

Índice general

Introducción	v
1. Nociones de Optimización Robusta	1
1.1. Introducción	1
1.1.1. Robustez y Optimización	2
1.2. Revisión Bibliográfica	5
1.2.1. Análisis de Sensibilidad	5
1.2.2. Resolución de Problemas Robustos	6
1.2.3. Comentarios sobre la Geometría de los Problemas	8
1.3. Definiciones Generales	10
1.3.1. Ejemplo en 2D	11
2. Sensibilidad del Valor Óptimo a Incertidumbre Poliedral	12
2.1. Introducción	12
2.2. Definiciones y Notación	14
2.3. Modelo para la Estimación de la Sensibilidad del Valor Óptimo de un Problema Lineal con Incertidumbre Poliedral de sus Datos	21
2.3.1. Construcción de la Estimación	22
2.3.2. Modelamiento de la Estimación	25
2.3.3. Cálculo de la Estimación	28
3. Dos Caracterizaciones de Incertidumbre Poliedral	33

ÍNDICE GENERAL

3.1. Introducción	33
3.2. Dos Modelos de Sensibilidad con Incertidumbre Poliedral	36
3.2.1. Modelo Bajo Incertidumbre Independiente en los Coeficientes	36
3.2.2. Modelo Bajo Incertidumbre Independiente en las Restricciones	38
3.3. Comentarios para el Modelador	39
4. Estimación de la Sensibilidad de Problemas de Reales en NETLib	40
4.1. Introducción	40
4.2. Aplicación de Nuestra Estimación	41
4.2.1. Librería de Problemas	41
4.2.2. Algoritmo	42
4.2.3. Resultados Obtenidos	43
5. Análisis y Conclusiones	48
5.1. Introducción	48
5.2. Análisis de los Resultados	48
5.2.1. Conclusiones	50
5.3. Trabajo Futuro	52
5.3.1. Tiempo de Ejecución	52
5.3.2. Nivel de Incerteza y Restricciones Redundantes	53
5.3.3. Solución Robusta	54
Apéndice	55
Anexo A	55
Anexo B	56
Anexo C	58
Anexo D	59
Bibliografía	63

Introducción

En el siglo dieciocho, la palabra *robusto* era utilizada para referirse a alguien que era fuerte, pero vulgar y violento a la vez. A mediados del siglo diecinueve la palabra comenzó a perder su connotación negativa cuando se comenzó a utilizar en modelación matemática. Hoy, *robusto* simplemente quiere decir fuerte, saludable o suficientemente resistente para superar las adversidades de la vida. En este mismo sentido, la Optimización Robusta propone métodos de optimización que son capaces de encontrar soluciones resistentes a las posibles variaciones de la información provista.

En general, un problema de decisión está sometido a una gran cantidad de factores que pueden provocar incertidumbre en la información, haciendo que las respuestas del modelo no sean del todo confiables. Existen en la literatura diversos métodos para hacerse cargo de la incerteza en los datos, como el análisis de escenarios, la optimización estocástica y la optimización robusta. En particular el problema de optimización robusta busca una solución que sea factible bajo cualquier escenario posible a partir de las eventuales variaciones en los valores de los parámetros, reflejadas en un *Modelo de Incertidumbre*. Resulta interesante ser capaz de predecir cuánto “arriesgamos”, en términos de la función objetivo, al utilizar una solución óptima que ignora la incertidumbre en lugar de una solución robusta. Ben-Tal y Nemirovski (1998) proponen el *Margen de Factibilidad* como una medida capaz de acotar dicho valor en términos de la geometría del problema y el nivel de incerteza al que se someten sus datos.

La Optimización Robusta en general, tiene una larga y nutrida historia, es objeto del presente trabajo examinar parte de esa historia y su relación con el estudio de investigación realizado. La investigación en esta tesis se preocupa de encontrar una medida computacionalmente abordable de la sensibilidad de la función objetivo de problemas lineales frente a datos con variaciones cuyo modelo de incertidumbre es poliedral. Esto permitirá construir un test que evaluará la utilidad (o bien el costo) de la estrategia conservadora propuesta por Ben-Tal y Nemirovski para problemas de optimización con incertidumbre, con la intención de brindar información *a priori* al momento de abordar la resolución computacional de estos problemas.

Esta tesis se estructura de la siguiente forma. En el Capítulo 1 se introduce al lector en el mundo de la Optimización Robusta haciendo una revisión bibliográfica desde los comienzos de esta teoría abordando en detalle el Análisis de Sensibilidad y la resolución de problemas de Optimización a través de estrategias conservadoras o Algoritmos Robustos, además se introducen las definiciones y

notaciones necesarias para el desarrollo de la investigación.

En el Capítulo 2 se elabora, a partir de una primera versión provista por Ben-Tal y Nemirovski, la teoría que permite construir la estimación que se busca para un modelo de incertidumbre generalizado, y descrito por un conjunto poliedral. La estimación propuesta es finalmente, producto de mejoras y adaptaciones de la versión original. Dado que la estimación se contruye para ser utilizada bajo cualquier conjunto de incerteza poliedral, el Capítulo 3 tiene como finalidad exponer dos modelos particulares de incertidumbre en los cuales la metodología construida puede ser aplicada.

Ben-Tal y Nemirovski, en una segunda publicación (2000), escogieron un modelo de incertidumbre para evaluar la sensibilidad de los problemas de la librería NETLib. En el Capítulo 4 se bosqueja el algoritmo implementado para calcular la estimación para el mismo modelo utilizado por los autores señalados y se detallan los resultados obtenidos para el mismo grupo de problemas de control.

El análisis de los resultados, comparaciones con estudios similiares y todas las conclusiones que se obtuvieron en el desarrollo de esta investigación se resumen en el Capítulo 5 donde además se plantean algunas tareas pendientes y se enuncian algunas ideas para desarrollos futuros en esta misma índole. Finalmente, se incluye un apéndice con información adicional.

Esta investigación constituye el proyecto de título de Rodrigo López Insinilla, estudiante del Programa de Magister en Gestión de Operaciones de la Universidad de Chile e Ingeniería Civil, mención Matemática. Este trabajo se realizó con la vital colaboración de Fernando Ordoñez, Roberto Cominetti, Daniel Espinoza y Jorge Vera, académicos de la Universidad de Chile y Pontificia Universidad Católica de Chile, todos miembros de la comisión evaluadora.

Este trabajo se originó a partir de una pasantía académica que el alumno realizó en el *Viterbi School of Engineering* de la *University of Southern California* durante el año 2008, en el marco del Programa de Movilidad Estudiantil para Estudiantes de Postgrado de la Universidad de Chile. Contó además con el apoyo financiero del Instituto Milenio de Sistemas Complejos de Ingeniería y el Departamento de Ingeniería Matemática de la Universidad de Chile.

Capítulo 1

Nociones de Optimización Robusta

Los avances tecnológicos permiten enfrentarnos a problemas cada vez más complejos: La creciente demanda por soluciones que se ajusten a la realidad introduce un complejo desafío ¿Cómo hacerse cargo de la entropía presente en la naturaleza para ser capaces de generar soluciones con buen comportamiento frente a todas las eventualidades?

1.1. Introducción

Cada vez que tomamos una decisión, observamos *todos* los factores *relevantes* y evaluamos las *posibles alternativas*. Este simple acto solamente es correcto al asumir que los factores que determinan las posibilidades son conocidos de forma certera, pero la verdad es que en la vida diaria es difícil encontrar factores que sean fijos o estables. En general asumimos que esta variabilidad no puede afectar en gran manera nuestra decisión, o bien estamos dispuestos a aceptar las consecuencias que una decisión errónea puede tener, pero en los complejos sistemas de la Ingeniería una mala decisión puede llegar a ser muy costosa o generar errores irreparables, por lo que no es recomendable permitirnos ignorar la variabilidad de los factores del sistema.

Para comenzar a construir métodos matemáticos que logren hacerse cargo de la incertidumbre de la información, primero debemos acordar qué es lo que entendemos como incerteza, y cómo esta, puede llegar a afectar al modelo matemático que describe o resuelve un sistema. La incerteza, por definición, es la inseguridad frente a la veracidad de un hecho. En un modelo matemático se entiende como la imposibilidad de aceptar como cierto el valor de los datos provistos. Naturalmente la respuesta de cualquier modelo depende exclusivamente de los datos que este recibe, de modo que si uno o más parámetros son inciertos, esto se traduce en una respuesta poco confiable. Los modelos de Optimización Matemática han sido importantes desde el siglo XX, pues son una valiosa herramienta para el apoyo en la teoría de toma de decisiones (complejas decisiones), por lo que se hace necesario contar con un método que permita hacerlos aplicables a sistemas modernos que incorporan datos inciertos.

En un ambiente incierto, la incertidumbre puede ser producto de un efecto combinado de varios motivos, puede ser de muy diversa naturaleza, e incluso puede estar presente sólo durante algunos periodos de tiempo. Independiente de cómo esta se manifieste, afectará al ambiente de decisión poniendo en riesgo la optimalidad y/o factibilidad de las soluciones obtenidas. Frente a esto existen varias estrategias que permiten encontrar soluciones confiables. Para comenzar a enfrentar este problema, es clave entender el comportamiento de la variabilidad. Un dato incierto se *mueve*, tal vez, respondiendo a una distribución de probabilidad. Si tuviésemos la suerte de conocer aquella distribución, o si pudiésemos estimarla, entonces es posible encontrar soluciones que tienen *alta probabilidad de ser óptimas* a través de metodologías como la Optimización Estocástica, ver [7], [22], [18] y [24].

Por otro lado, conocer una estimación de esta distribución nos permitiría realizar estudios con herramientas de Simulación (generación de instancias posibles) lo que nos llevaría a encontrar soluciones que son *factibles para una gran cantidad de escenarios*, ver [25], [1], [9], [12] y [13]. A pesar que estas herramientas son útiles sólo cuando la aleatoriedad de los datos es de naturaleza estocástica, y más aún, cuando es posible conocer o estimar su distribución de probabilidad, en general han resultado ser muy efectivas en la práctica y su uso es muy popular en la ingeniería.

Cuando es imperativo que el tomador de decisiones sea adverso al riesgo, por ejemplo cuando sus decisiones pueden tener consecuencias catastróficas, o cuando la decisión debe tomarse para periodos prolongados de tiempo, o en cualquier situación donde es preferible (y posible) *invertir recursos* con el fin de seguir una estrategia óptima confiable (que incluya la aversión al riesgo del tomador de decisiones), es necesario considerar una estrategia factible en todos los escenarios posibles generados por la incertidumbre. De todo esto y más, la Optimización Robusta se hace cargo, manteniéndose a la vanguardia de las herramientas que permiten incluir un espíritu conservador en la resolución de problemas de una forma computacionalmente abordable.

Como es natural, optimizar considerando todas las regiones factibles posibles a la vez, es un problema complejo, que en la mayoría de los casos se puede aproximar a través de un problema semi-definido, por lo que en ocasiones, antes de invertir recursos computacionales en la resolución de un problema de esta naturaleza, es deseable poder conocer cuánto podría llegar a variar la respuesta en virtud de la incerteza presente¹.

1.1.1. Robustez y Optimización

La Real Academia Española define el término *robusto* como *fuerte o de buena salud*. Los modelos matemáticos modernos han demostrado ser capaces de integrar esta noción, al mismo tiempo que logran extender sus aplicaciones a sistemas aún más complejos. La Optimización Robusta es un método o estrategia de optimización que busca incorporar la entropía del mundo real como parte del problema, es decir, pretende hacerse cargo de la variabilidad de los parámetros del problema con el fin de lograr una modelación resistente a los eventuales cambios, y que entregue respuestas que

¹Esto es lo que entendemos como *Sensibilidad*.

sean inmunes a estos.

La necesidad de contar con una herramienta tecnológica que apoyara la toma de decisiones se remonta a los años 50, cuando se trazaban las primeras líneas de la teoría moderna de toma de decisiones². Con el paso del tiempo, fueron necesarios modelos más *realistas*, capaces de incorporar objetivos y restricciones que se ajustaran al dinamismo de los problemas fuera de la pizarra. A partir de esto, la modelación de sistemas y sus herramientas de resolución comenzaron a evolucionar permitiendo incluir *la aleatoriedad de la información*. Las primeras investigaciones académicas que buscan este tipo de modelamiento se publicaron casi veinte años después, ver [17] y [23].

Con el fin de estudiar las nociones que permiten construir una estrategia conservadora, detallaremos a continuación la construcción de un modelo generalizado de optimización robusta. Sea \mathcal{D} un conjunto de escenarios posibles y sea un vector de datos $\zeta \in \mathcal{D}$. Entendemos como *Optimización* al modelamiento matemático de tres elementos: un conjunto de **variables de decisión**, $x \in \mathbb{R}^n$, cuyo valor debe determinarse, un **objetivo**, $f(x, \zeta) \in \mathbb{R}$, que buscamos llevar a un nivel ideal a través de estas variables, y un conjunto de **restricciones**, $F(x, \zeta)$, que describen el conjunto de posibilidades para las variables de decisión. La Optimización Robusta incorpora al modelamiento un cuarto elemento: la estructura del conjunto \mathcal{D} , o bien el equivalente matemático de la incerteza de los datos los datos que definen el objetivo y las restricciones en ζ . A esto se le llama **Modelo de Incerteza**.

Para un valor fijo o *nominal* de datos ζ_0 consideremos la conceptualización más clásica para esta modelación, dada por el problema (P).

$$(P) \begin{cases} \text{mín} & v \\ \text{s.a.} & \\ & v - f(x, \zeta_0) \geq 0 \\ & F(x, \zeta_0) \geq 0 \end{cases}$$

Un vector de variables de decisión x factible para todas las instancias de la familia de problemas posibles debe cumplir

$$v - f(x, \zeta) \geq 0 \quad \forall \zeta \in \mathcal{D} \tag{1.1}$$

$$F(x, \zeta) \geq 0 \quad \forall \zeta \in \mathcal{D} \tag{1.2}$$

Cumplir (1.2) asegura la factibilidad inmune a la variabilidad de los datos. Observemos que a partir de (1.1) valor del vector $x \in \mathbb{R}^n$ en el óptimo corresponde al valor del peor caso. Considerando (1.1) y (1.2), construimos un nuevo problema de optimización que entrega un óptimo inmune o robusto. A este problema lo denominamos **Contraparte Robusta** y para construirlo es necesario consensuar lo que entendemos por incertidumbre matemática.

A través de la literatura (ver [3], [6], [23] y [19]), podemos observar que el Modelo de Incerteza puede

²Recordemos que el Algoritmo Simplex [10] fue inventado por Dantzig a comienzos de los 50, cuando se requería la resolución de problemas lineales para tomar decisiones estratégicas en la logística militar post Segunda Guerra Mundial.

expresarse en varias formas, pero tiene elementos comunes que permiten sostener una estructura tipo que permita incorporar la incerteza al modelo de optimización. Estos elementos son:

- El **Impacto de la Incerteza** sobre el problema. Dice relación a la cantidad o al conjunto de parámetros que se ven afectados por la incertidumbre del ambiente de decisión. Existen varias razones para aceptar en el modelamiento valores poco confiables para algunos datos, ya sea porque este valor es desconocido al momento en que debe tomarse una decisión (en cuyo caso se estila usar una estimación), o porque representa un valor que no puede ser medido, estimado o calculado con precisión producto de su naturaleza. A estos coeficientes los denominamos *incierto*s. Así mismo, existirán parámetros que son fijos o estructurales, en otras palabras, confiables, por lo que la incerteza no los afectará.
- El **Tamaño de la Incerteza**. Corresponde al nivel o grado en que los parámetros son afectados, es decir, permite cuantificar o controlar el nivel de confiabilidad que tenemos sobre los datos. La variabilidad puede ser expresada de muchas formas, las más usuales son la descripción de un conjunto posible de datos, la variación porcentual de los valores respecto a una versión nominal o estimación, o también se suele convenir una cantidad de dígitos o decimales que deben coincidir con alguna versión nominal. Ver [4] y [23].
- La **Correlación de la Variabilidad**. Permite entender el efecto combinado que puede darse en grupos de parámetros inciertos. A través de esta noción es posible entender que algunos parámetros varíen sólo cuando otros lo hacen o no, o bien que el crecimiento o decrecimiento de algunos parámetros sea común o contrario entre parámetros. Este elemento lo podemos cuantificar a través de la matriz de covarianza de los parámetros.
- El **Comportamiento de la Variabilidad**. La variabilidad de los datos por sí sola puede medirse y estudiarse, existen tres posibilidades: Naturalmente asumir que los datos pueden seguir alguna distribución de probabilidades. Si no fuese así, o bien estarán acotados por un valor máximo y mínimo, o en el peor de los casos, tendrán variabilidad irrestricta. Esta característica de la incertidumbre resulta ser muy importante para el Modelo de Incertidumbre que se escoja, pues estas tres alternativas definen los caminos en los que la Optimización Robusta se ha desarrollado hasta ahora.

Consideremos el problema (P). Cuando los parámetros son afectados por la incertidumbre, el modelo de incerteza se activa y se genera una familia de problemas posibles. Para cada $\zeta \in \mathcal{D}$ los problemas tendrán la forma

$$(P_\zeta) \left\{ \begin{array}{l} \text{mín} \quad v \\ \text{s.a.} \\ v - f(x, \zeta) \geq 0 \\ F(x, \zeta) \geq 0 \end{array} \right.$$

De este modo, la *Contraparte Robusta* que se obtiene esta dada por

$$(P^*) \begin{cases} \text{mín} & \sup_{\zeta \in \mathcal{D}} f(x, \zeta) \\ \text{s.a.} & F(x, \zeta) \geq 0 \quad \forall \zeta \in \mathcal{D} \end{cases}$$

El estudio y resolución de este problema ha tenido una interesante evolución en los últimos cuarenta años. A continuación haremos una revisión de algunos de los resultados más importantes.

1.2. Revisión Bibliográfica

Antes de revisar los avances más importantes que se conocen en Optimización Robusta, es importante que aclaremos el uso o abuso de este concepto. Hasta ahora, nosotros hemos llamado Optimización Robusta, en términos generales, a la teoría matemática que incorpora dentro de sus elementos la incerteza de los datos. Ahora bien, cualquier herramienta que genere soluciones inmunes a la incertidumbre de los parámetros, la denominaremos por *robusta* (al igual que sus soluciones), pero también dentro de esta teoría matemática se consideran algunas metodologías cuyo fin no es generar soluciones robustas, pero si estudiar este fenómeno en términos de la sensibilidad o comportamiento de los elementos del problema bajo ambientes inciertos. Se aclara entonces, que la *Modelación u Optimización Robusta* y el *Análisis de Sensibilidad* son dos conceptos diferentes. El primero busca proteger a la solución de futuros cambios en los datos, mientras que el segundo busca cuantificar el efecto de las variaciones de los datos sobre la solución y/o el valor óptimo. En este sentido, una solución (completamente) robusta, **no** es sensible a variaciones de los parámetros.

1.2.1. Análisis de Sensibilidad

Al Análisis de Sensibilidad se le denomina una práctica *Post Mortem*, puesto que el estudio se realiza posterior a la ejecución de algún algoritmo, en general estas estrategias no son populares puesto que las características de la solución se conocen una vez que ya se han invertido importantes recursos. Nuestra investigación busca innovar en cuanto a lo que se conoce en Análisis de Sensibilidad, proponiendo un método de cálculo a priori de la **sensibilidad de la función objetivo** frente a **restricciones inciertas** de problemas lineales. La ventaja de nuestro método es que no necesita la resolución del problema original para realizar la estimación.

La versión clásica del Análisis de Sensibilidad para Optimización Lineal es el denominado Análisis Post-Optimal, que consiste en resolver una instancia nominal del problema, para estudiar el rango de variación que admiten sus datos para que la base óptima siga siendo base, en otras palabras, para que la solución óptima siga siendo factible y se mantenga óptima. El mayor defecto de esta estrategia radica en el hecho que en la mayoría de los casos, la solución sólo sigue siendo óptima para variaciones muy pequeñas, y no es posible entregar información respecto a escenarios con variaciones mayores. Más aún, solo entrega información de la solución nominal, la cual no necesariamente es

representativa de los otros escenarios posibles.

Otra alternativa es el *Análisis de Escenarios*, que consiste en resolver un grupo considerable de escenarios posibles, con el fin de calcular la variación que podría experimentar la función objetivo. A partir de estos resultados se puede tener una buena noción de cómo se ve afectado el valor óptimo producto de la incerteza, pero no se puede garantizar que la solución de cada escenario sea factible en los otros, por lo que la factibilidad no resulta ser inmune y no existe criterio alguno para elegir un escenario como representativo. Otro defecto en esta estrategia radica en que se requiere la resolución de una cantidad importante de problemas nominales, lo que en ocasiones la hace poco abordable computacionalmente.

Nuestra propuesta permite estimar la sensibilidad del valor óptimo, a partir de una única versión nominal, considerando un modelo de incerteza para problemas lineales bastante general. El fin de nuestro método es evaluar la ventaja (o desventaja) de la costosa resolución de un problema robusto, y en caso que se opte por él, entrega un punto robusto factible, que facilita la posterior ejecución del algoritmo.

1.2.2. Resolución de Problemas Robustos

La Optimización Robusta comenzó con el uso del *Análisis del Peor Caso* y el *Modelo Maximin de Wald* como herramientas para decisiones bajo ambientes inciertos, sin embargo estas no eran lo suficientemente versátiles. En 1955, Dantzig pone la primera piedra para el tratamiento de la incerteza [11], adaptando el algoritmo Simplex para Problemas Lineales con Datos Inciertos. La robustez se convirtió en un campo de estudio como tal en los años 70 cuando fue necesario obtener soluciones robustas en áreas específicas de la Ingeniería como la Investigación de Operaciones, la Teoría del Control, la Economía, entre otras. En una primera instancia se realizaron desarrollos en paralelo para las necesidades específicas de estas áreas, ver por ejemplo [16] y [8]. Ya en los 90 y a comienzos de este siglo comenzaron las publicaciones de avances más amplios, ver [5], [3], [6], [2], [21], [14] y [15].

En cuanto a generación de soluciones inmunes, es bastante lo que se ha avanzado. El caso más general corresponde a asumir que los datos simplemente varían sin responder a ninguna ley. En este caso, hay dos corrientes, que aunque están de la mano generan estudios independientes. Se dará que, o bien los datos están acotados, o bien no hay restricción alguna sobre ellos. El primer caso dice relación a situaciones donde la variación está restringida a un intervalo conocido, y el segundo a casos en que no se conoce mayor información de los datos o simplemente, cuando se desconoce su valor. Como ya lo señalamos, si los datos varían respondiendo a una distribución de probabilidades, entonces se pueden aplicar otro tipo de herramientas.

Aún sin utilizar el término *robusto*, avances preliminares en esta área fueron tomados por Gould [17] quien modificó el Algoritmo Simplex para encontrar óptimos inmunes a la inexactitud del vector de

recursos en Problemas Lineales, pero con tolerancia conocida, es decir, para problemas de la forma

$$\begin{cases} \text{mín} & c^T x \\ \text{s.a.} & Ax \leq b \pm e \\ & x \geq 0 \end{cases}$$

En el Anexo A se incluye un bosquejo construido por Gould en [17] donde se aprecia cómo puede variar la región factible en un problema de dos variables cuando su vector de recursos sufre variaciones, e incluso como restricciones que eran redundantes, pueden llegar a ser activas.

Poco después, Soyster en [23] propone un modelo convexo para resolver un problema de Programación Lineal Inexacta. Su modelo de incerteza es bastante natural, supongamos un problema de la forma

$$\text{mín } c^T x \quad \text{s.a.} \quad a_1 x_1 + a_2 x_2 + \cdots + a_n x_n \leq b \quad ; \quad x \geq 0$$

donde cada $a_i \in K_i$ un conjunto convexo de valores posibles para el coeficiente incierto. Como el óptimo robusto debe satisfacer la restricción para cualquier valor posible de sus coeficientes, y como x tiene todas sus coordenadas positivas, es directo que el caso más desfavorable es aquel en que cada coeficiente toma su valor más alto, es decir, para encontrar el óptimo robusto basta resolver

$$\text{mín } c^T x \quad \text{s.a.} \quad (\max_{a \in K_1} a)x_1 + (\max_{a \in K_2} a)x_2 + \cdots + (\max_{a \in K_n} a)x_n \leq b \quad ; \quad x \geq 0$$

Tuvieron que pasar 20 años para que el paradigma de la robustez evolucionara más profundamente. Hasta ese entonces, los métodos a pesar de ser completamente confiables, parecían ser en extremo conservadores, tal vez demasiado para las aplicaciones en la Ingeniería ¿*Por qué, en caso de variar los datos, estos deben hacerlo de la peor forma posible?* o ¿*Por qué deben hacerlo todos simultáneamente?* Recién a comienzos de los años 90 el paradigma de la robustez comienza a ser lo que hoy es, cuando se sugiere la posibilidad de explorar soluciones que no sean *completamente* robustas y que se obtienen usando aproximaciones de la región robusto factible. De esto, nace la tendencia a considerar modelos menos conservadores para conseguir soluciones que son inmunes sólo **en la mayoría de los casos**, o bien, soluciones que tienen una alta probabilidad de ser robusto factibles.

Ben-Tal y Nemirovski fueron precursores en esta idea, en [3] definieron una contraparte robusta que utiliza un conjunto elipsoidal para aproximar la región factible robusta, por supuesto, esta aproximación tiene un costo que se refleja en las propiedades del problema a resolver, por ejemplo el caso lineal tiene una contraparte robusta cuadrática. Existe basta evidencia (ver [5] y [2]) que en la práctica es posible aceptar una aproximación de este estilo.

Por otro lado, Bertsimas y Sim en [6] también proponen una alternativa de aproximación. Usando una contraparte lineal, son capaces de regular el grado de conservadurismo de las soluciones, a través del número de coeficientes inciertos (Impacto de la Incerteza). Ellos proponen que una restricción con muchos coeficientes inciertos tiene mayor probabilidad a ser violada que una con pocos. De este modo, se construye una región de vectores con alta probabilidad de ser robustos y que en cierto

modo aproxima la región factible robusta.

Los últimos avances [26] muestran que es posible generar contrapartes robustas convexas para problemas lineales sin la necesidad de generar demasiadas variables auxiliares. Un paso adelante respecto a la complejidad computacional de las estrategias robustas, pero aún nos enfrenta a algoritmos no lineales. En general, aproximaciones de este tipo requieren vasto recurso computacional para resolverse, lo que hace que optar por una alternativa robusta sea una decisión importante, y por ende un análisis de sensibilidad previo es una buena herramienta de análisis para un tomador de decisiones.

1.2.3. Comentarios sobre la Geometría de los Problemas

Hasta ahora simplemente hemos aceptado que la incerteza de los datos afecta de una u otra forma las respuestas que cualquier modelo matemático puede ofrecer, y no hemos hecho mención alguna al porqué ocurre esto. *¿Qué es lo que hace que un dato incierto sea tan nocivo? ¿Por qué hay problemas que son más sensibles que otros?* Al menos el caso lineal, existe una explicación fácil de digerir: La geometría de los problemas.

Sabemos que las soluciones óptimas de problemas lineales se encuentran en alguna de las caras del poliedro factible. Además, cada una de ellas es potencialmente en extremo sensible a variaciones de los datos. En particular, el *movimiento* de los vértices, tendrá un efecto muy significativo en la sensibilidad del valor óptimo. En la Figura 1.1 es posible observar el impacto que la incerteza tiene en la geometría del problema.

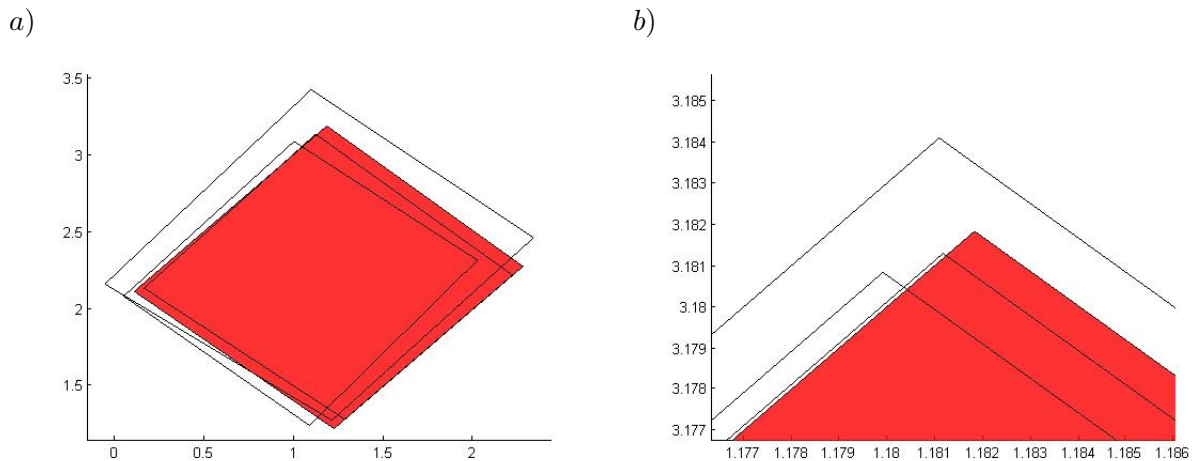


Figura 1.1: Poliedro de la forma $Ax \leq b$ y tres poliedros obtenidos con perturbaciones aleatorias de la forma $a(1+\varepsilon\delta)$ con $\delta \sim U(-1, 1)$ sobre todos los coeficientes de la matriz A original. **a)** Nivel de incerteza $\varepsilon = 0,1$. **b)** Detalle del vértice superior de las mismas perturbaciones, pero con $\varepsilon = 0,001$

Ahora bien, para encontrar la región robusto factible, debemos considerar todas las perturbaciones posibles. Al hacer esto, podemos ver en la Figura 1.2, que la geometría del problema sufre

considerables variaciones.

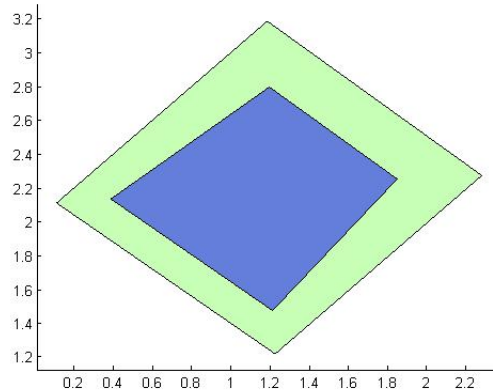


Figura 1.2: Poliedro de la Figura 1.1 y su región robusto factible considerando un nivel de incerteza del 10 %

Estas serán más o menos nocivas dependiendo de la geometría del problema. Consideremos un poliedro donde uno de sus vértices es un ángulo muy agudo, como él que se describe a continuación. (Ver Figura 1.3.a).

$$\begin{aligned} x + 40y &\leq 40 \\ x - 40y &\geq 40 \\ 2x + y &\geq 2 \end{aligned}$$

Uno de sus vértices es $(40 ; 0)$. Al considerar **una** variación aleatoria de los datos del 10 % se obtiene una traslación de este vértice a $(36,1 ; 0)$. De modo que cualquier función objetivo cuya primera variable sea importante será fuertemente sensible. (Ver Figura 1.3.b).

¿Cuál es la característica geométrica que hace a algunos problemas más sensibles que otros? La Teoría del Número de Condicionamiento (de Problemas Convexos) ha evidenciado que dicha medida está relacionada con una gran cantidad de propiedades geométricas y algorítmicas de problemas de optimización, en particular nos interesa su relación con la deformación que sufren las regiones factibles y la variación del valor óptimo a partir de variaciones en los datos. En [20] se da una completa revisión bibliográfica de esta relación, y se calcula el número de condicionamiento para el mismo grupo de problemas lineales que usaremos como grupo de control. Esto nos da la oportunidad de respaldar nuestros cálculos con otra mirada del mismo problema, al mismo tiempo que nos permitirá respaldar la hipótesis que el número de condicionamiento explica el tamaño del impacto que tiene la incerteza sobre la geometría, o dicho de otra manera, el nivel de sensibilidad de los problemas. Complementaremos nuestro estudio con los cálculos que sugiere esta teoría. Ver Capítulo 4.

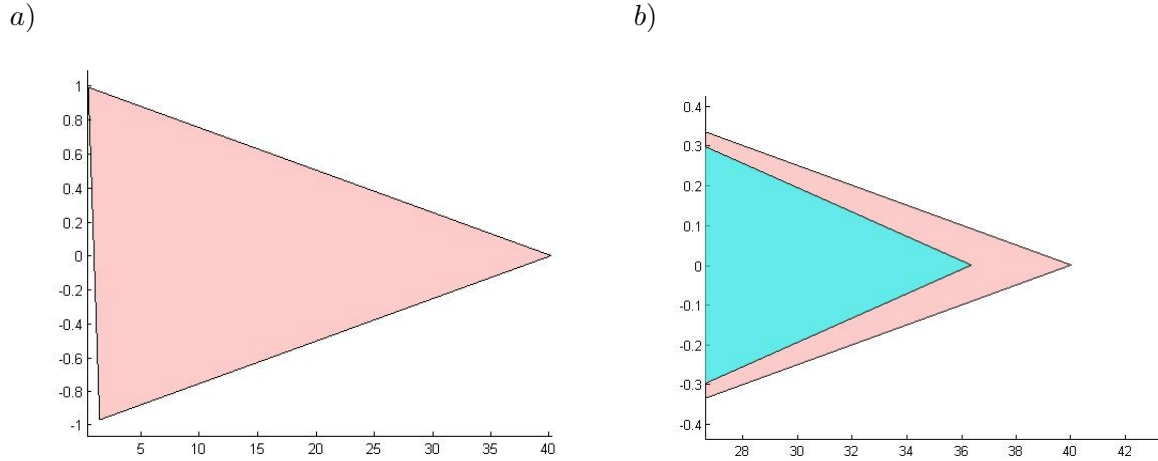


Figura 1.3: **a)** Poliedro $Ax \leq b$. **b)** Detalle del vértice derecho de una perturbación aleatoria sobre los coeficientes de la matriz original con nivel de incerteza $\varepsilon = 0,1$. *Nota: Los ejes están rescalados*

1.3. Definiciones Generales

Para abordar nuestra investigación, no es necesario tener mayores conocimientos en Optimización Robusta, pero si es indispensable manejar la terminología utilizada, por lo que hemos incluido esta sección a modo de glosario.

DEFINICIÓN 1.1. Problema Incierto. Llamamos *Problema Incierto* a aquel problema de optimización en el que los valores conocidos de sus parámetros no son conocidos con exactitud. La incertidumbre afecta solo al valor de sus datos, es decir, asumimos que la estructura del problema (variables y restricciones) es fija, y confiable. Notaremos un problema incierto como $(\mathcal{P}_\varepsilon)$ donde ε denota el nivel de incerteza (máximo) de sus datos.

DEFINICIÓN 1.2. Modelo de Incerteza. Cuando hablamos de *Modelo de Incerteza* en general, hacemos referencia a los cuatro elementos que lo conforman: *Impacto en el Problema, Tamaño de la Incerteza, Correlación de la Variabilidad y Comportamiento de la Variabilidad.*

DEFINICIÓN 1.3. Conjunto de Incerteza. El *Conjunto de Incerteza* es el conjunto que nos permite describir todos los posibles valores que pueden tomar los parámetros del problema. En general, cuando se habla de *Incetza Poliedral* o *Incetza Elipsoidal*, se hace haciendo referencia a la geometría de este conjunto. Lo denotaremos por \mathcal{U} .

DEFINICIÓN 1.4. Coeficientes Inciertos La incerteza del problema puede afectar a un subconjunto de coeficientes de A , o bien a todos a la vez. Notaremos por I al conjunto de todos los coeficientes inciertos (por simplicidad notaremos $(i, j) \in I$ como $ij \in I$).

DEFINICIÓN 1.5. Problema Robusto. (o *Contraparte Robusta*) Es aquel problema de optimización que permite encontrar soluciones que son inmunes a la incertidumbre. El problema robusto

contiene los elementos del problema original e incorpora la incertidumbre a través de las convenciones tomadas para el modelo de incerteza. La modelación estándar para la contraparte robusta considera todas las restricciones posibles, en otras palabras, la región factible será la intersección de todas las regiones factibles posibles. Esto generará un problema semidefinido pero que puede ser desarrollado para conseguir mejores formulaciones.

DEFINICIÓN 1.6. Región y Solución Robusta. Corresponden a la región factible y a la solución del problema robusto respectivamente. Los denotaremos por \mathcal{F} y x^* .

1.3.1. Ejemplo en 2D

A continuación se detalla el poliedro que nos acompañará, a modo de ejemplo, en los siguientes capítulos.

$$\begin{aligned} -x + y &\geq 0 \\ x - y &\geq -2 \\ x + 1,2y &\leq 5 \\ 0,8x + y &\geq 2,2 \end{aligned}$$

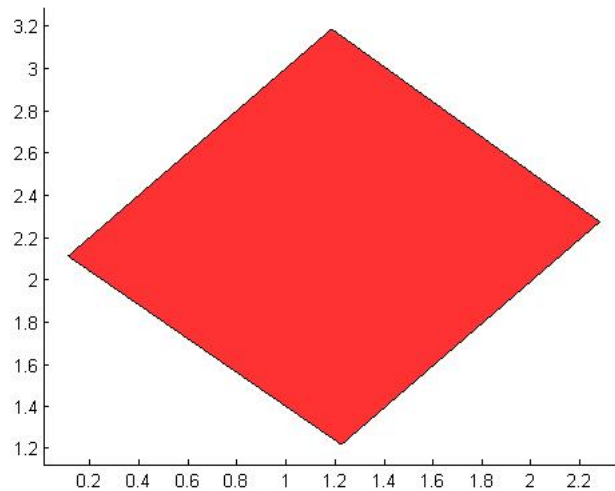


Figura 1.4: Poliedro en \mathbb{R}^2 .

Capítulo 2

Sensibilidad del Valor Óptimo a Incertidumbre Poliedral

Para un modelador no es poco usual enfrentarse a escenarios donde la información es desconocida o poco confiable. La Optimización Robusta se hace cargo de esta incerteza entregando soluciones *confiables*, pero ¿Cuánto cuesta entregar esta confianza?

2.1. Introducción

Como mencionamos en el capítulo anterior, es posible aplicar estrategias matemáticas que permiten incorporar dentro de un modelo lo incierto que pudiesen ser algunos de sus datos. Ya sea producto de la naturaleza del coeficiente o por su total desconocimiento, esta *volatilidad* del parámetro se puede modelar para dar paso a algoritmos que encontrarán una *mejor* solución que se mantenga factible **a pesar de todas** las posibles eventualidades, es decir, que independiente del valor que tomen los coeficientes inciertos, esta solución siempre será factible. Dichos algoritmos son propuestos por la Optimización Robusta, pero hacerse cargo de la incerteza de los datos no es matemáticamente fácil de abordar, pues es necesario asumir supuestos que harán del problema un asunto complejo. Esta complejidad aumentará el tiempo de resolución del problema (independiente de la capacidad del procesador que se utilice), por lo que no es trivial decidir el uso de una estrategia conservadora sin saber con anticipación si esta entregará resultados prácticos.

La sensibilidad de un problema frente a la incerteza nos permite *cuantificar* el valor de una estrategia conservadora aplicada a dicho problema. A continuación revisaremos los elementos que permiten construir una estimación de la sensibilidad del valor óptimo de un problema de optimización lineal. Visitaremos las hipótesis necesarias para construir un modelo matemático aplicable y comparable. Esta estimación será una herramienta *a priori*, pues nos permitirá evaluar, previo a la aplicación de un costoso algoritmo robusto, el impacto que tendría una postura conservadora en el valor óptimo del problema, en relación a su versión nominal.

Consideremos un problema de optimización con función objetivo sin incertidumbre $f(x)$ y con vector de datos (inciertos) ζ . Supongamos que conocemos alguna versión de sus coeficientes inciertos dada por ζ^0 . Sea x^0 el óptimo de esta instancia del problema y x^* el óptimo robusto (calculado con algún algoritmo proporcionado por la Optimización Robusta). El valor $|f(x^*) - f(x^0)|$ se puede entender como la *sensibilidad* del valor óptimo del problema nominal. Si este valor es pequeño, entonces el óptimo nominal es una solución que podemos considerar como *acceptable* y de este modo no sería necesario invertir los grandes recursos de procesamiento que se consumen en la aplicación de un algoritmo robusto. En cambio, si esta distancia es significativa, entonces la inversión si es necesaria. Para un tomador de decisiones que enfrenta un problema de optimización de gran tamaño, le resultará de gran utilidad poseer una herramienta que testeé, sin consumir muchos recursos, si la aplicación de un algoritmo robusto merece el esfuerzo, o bien que le permita cuantificar cuan *lejos* estará su solución nominal de un óptimo conservador.

Nuestra estimación pretende entregar dicha información, trabajaremos en pos de conseguir una herramienta de cálculo aplicable a problemas reales. Nuestra intención será conseguir resultados exactos, si la herramienta resultase ser de utilidad entonces quedará pendiente la adaptación del algoritmo para encontrar formas de cálculo (o aproximación) eficientes.

Como revisamos en el capítulo anterior, es posible encontrar vasta literatura respecto a estrategias conservadoras de resolución de problemas de optimización. Si bien, nuestro modelo no requiere la resolución del problema robusto, ni tampoco conocer en detalle el algoritmo que busca este óptimo, si es necesario utilizar un modelo de incertidumbre, es decir, una expresión matemática que nos permita modelar el efecto de la volatilidad de los datos. Nuestra estimación dependerá directamente del modelo elegido, por lo que no somos indiferentes a la estrategia de optimización robusta elegida. Para la elección de esta, tendremos en cuenta elementos que sean más bien prácticos para el desarrollo de esta investigación:

- Las hipótesis del modelo deben permitir que el cálculo de la estimación sea abordable computacionalmente.
- El modelo debe ser aplicable a alguna biblioteca de problemas disponible, con el fin de utilizarla como grupo de control.
- El modelo debe ser comparable con otros estudios similares, ver [4] y [20].

Con el fin de lograr un modelo-algoritmo cuya implementación sea razonable y práctica, nos hace sentido restringir el alcance de nuestra estimación a una familia de problemas con características apropiadas y que permitan cumplir todos estos requerimientos. La Optimización Lineal contiene hipótesis idóneas para estos. Los problemas lineales son un grupo importante dentro de la ingeniería, por lo que la elección parece ser la adecuada.

Al considerar problemas lineales podemos asumir que la incertidumbre sólo se expresa en los coeficientes de la matriz de restricciones. Si un problema presentara incertidumbre en su función objetivo

o en su vector de recursos no es difícil reescribir el problema y traspasar la incertidumbre a la matriz, por lo que siempre será posible asumir esta estructura.

El modelo de incertidumbre considerado en este trabajo es desarrollado por Ben-Tal y Nemirovsky en [3] de forma bastante general, brindándonos la oportunidad de simplificarlo de forma natural a problemas lineales.

2.2. Definiciones y Notación

Dados $A \in \mathcal{M}_{m_1 \times n}$, $E \in \mathcal{M}_{m_2 \times n}$, $c, l, u \in \mathbb{R}^n$, $b \in \mathbb{R}^{m_1}$ y $e \in \mathbb{R}^{m_2}$. Consideremos la estructura de un problema de minimización (\mathcal{P}) en \mathbb{R}^n dada por

$$(\mathcal{P}) \begin{cases} \text{mín} & c^T x \\ \text{s.a} & \\ & Ax \leq b \\ & Ex = e \\ & l \leq x \leq u \end{cases}$$

SUPUESTO 2.1. *Asumiremos que los datos inciertos están sólo en las restricciones de desigualdad.*

En [4] Ben-Tal y Nemirovski justifican este supuesto en lo siguiente:

- Si hubiese incerteza en los datos de las restricciones de igualdad, quien modele el problema debe considerar que el valor del lado izquierdo varia producto de la incerteza y por lo tanto para cada una de estas restricciones, el modelador debe considerar un rango para Ex con cotas superior e inferior de la forma $[e - \epsilon_1, e + \epsilon_2]$.
- En estudios realizados anteriormente sobre los problemas en la librería **NETLib** se evidencia que la incerteza en los datos de las restricciones $l \leq x \leq u$ no afecta esencialmente la calidad de la solución nominal, lo que nos permite asumir estos datos como certeros.

La estructura definida anteriormente corresponde a la de los problemas de la librería **NETLib**, que será la librería de problemas que usaremos como grupo de control. Ahora consideremos un problema incierto $(\mathcal{P}_\varepsilon)$.

$$(\mathcal{P}_\varepsilon) \begin{cases} \text{mín} & c^T x \\ \text{s.a.} & \\ & \bar{A}x \leq b \\ & Ex = e \\ & l \leq x \leq u \end{cases}$$

La notación \bar{A} para la matriz de restricciones de desigualdad representa una posible manifestación de los coeficientes de esta matriz, es decir, \bar{A} es alguna de las matrices posibles. Cada una de las

eventuales matrices la relacionaremos con un único elemento $\delta \in \mathcal{U}$. Para cada uno de los coeficientes inciertos a_{ij} habrá un coeficiente δ_{ij} . Si a_{ij} es certero, entonces $\delta_{ij} = 0$. De aquí que $\mathcal{U} \subseteq \mathcal{M}_{m_1 \times n}$.

OBSERVACIÓN. *El caso $|I| = m_1 \times n$ es general pues δ puede tener parámetros nulos y esto podrá ser aprovechado por el algoritmo¹. Nuestro modelo busca ser lo más general posible por lo que siempre asumiremos que todos los coeficientes de A son inciertos, sin embargo, con el fin de guardar este efecto para la construcción del algoritmo mantendremos la notación I para los coeficientes inciertos.*

Asumiremos que \mathcal{U} tiene forma poliedral puesto que esto nos permitirá aplicar el modelo a problemas en circunstancias generales. Con el fin de describir esta sección de forma generalizada, para $\{M^i\}_{i=1}^{m_1}$ y $\{g^i\}_{i=1}^{m_1}$ apropiados² usaremos

$$\mathcal{U} = \{\delta \in \mathcal{M}_{n \times m_1} / M^i \delta_i \leq g^i, i = 1, \dots, m_1\}$$

Supongamos que conocemos alguna versión (\mathcal{P}_0) del problema (\mathcal{P}) .

$$(\mathcal{P}_0) \left\{ \begin{array}{l} \text{mín} \quad c^T x \\ \text{s.a.} \\ Ax \leq b \\ Ex = e \\ l \leq x \leq u \end{array} \right.$$

Este problema, su matriz A , su región factible F^0 , su óptimo x^0 y el valor óptimo de la función objetivo, los denotaremos por *nominal*.

SUPUESTO 2.2. *Los coeficientes de \bar{A} son una perturbación de sus coeficientes correspondientes en la matriz A de la forma*

$$\bar{a}_{ij} = (1 + \varepsilon \delta_{ij}) a_{ij}$$

Para los efectos de este capítulo asumiremos el nivel de incerteza $\varepsilon > 0$ es fijo.

El parámetro $\varepsilon > 0$ representa el nivel de incerteza (Tamaño de la Incerteza) que afecta a los datos y $\delta \in \mathcal{U}$ controla la perturbación de los a_{ij} 's (Comportamiento de la Variabilidad). Convengamos que A_{ij} denotará una matriz de ceros salvo por el coeficiente (i, j) que vale a_{ij} . Para explicitar la dependencia de δ y ε podemos escribir \bar{A} como

$$\bar{A} = A^\varepsilon(\delta) := A + \underbrace{\varepsilon \sum_{ij \in I} \delta_{ij} A_{ij}}_{\text{incerteza}}$$

Entendemos A^ε como la suma del término nominal y un término que contiene la incerteza que define cada instancia de \bar{A}

¹Ver Definición 1.4 en página 10.

²Depende del \mathcal{U} considerado, se detalla en Supuesto 2.18 en página 29.

Reescribimos el problema incierto como

$$(\mathcal{P}_\varepsilon) \left[\begin{array}{l} \text{mín} \quad c^T x \\ \text{s.a.} \\ A^\varepsilon(\delta)x \leq b \\ Ex = e \\ l \leq x \leq u \end{array} \right]_{\delta \in \mathcal{U}}$$

Se tendrá que³

$$\mathcal{F} = \left\{ x \in \mathbb{R}^n / \begin{array}{l} A^\varepsilon(\delta)x \leq b \quad \forall \delta \in \mathcal{U}, \\ Ex = e, \\ l \leq x \leq u \end{array} \right\}, \quad x^* = \arg \text{mín}_{x \in \mathcal{F}} c^T x$$

NOTACIÓN. Para $\lambda \geq 0$ definimos F^λ por

$$F^\lambda := \left\{ x \in \mathbb{R}^n / \begin{array}{l} A^\varepsilon(\lambda\delta)x \leq b \quad \forall \delta \in \mathcal{U}, \\ Ex = e, \\ l \leq x \leq u \end{array} \right\}$$

Esta notación es consistente con F^0 y $F^1 = \mathcal{F}$.

Notemos que $A^{\lambda\varepsilon}(\delta) = A^\varepsilon(\lambda\delta)$ por lo que podemos entender a F^λ como una modificación del nivel de incerteza que define a \mathcal{F} . A medida que λ crece F^λ representa conjuntos cada vez más conservadores, o más pequeños debido a que la incerteza aumenta y menos puntos robusto factibles la toleran. Se tiene que

$$0 < \lambda_1 < \lambda_2 < 1 < \lambda_3 \implies F^{\lambda_3} \subset \mathcal{F} \subset F^{\lambda_2} \subset F^{\lambda_1} \subset F^0$$

- $\lambda = 0$ describe un problema sin incerteza (problema nominal),
- $\lambda = 1$ describe un problema con la totalidad de la incerteza considerada (problema robusto),
- los valores de $\lambda \in (0, 1)$ describen problemas *intermedios*, o *en algún grado* conservadores, y
- los valores de $\lambda > 1$ describen problemas aún más estrictos que el problema robusto. Notar que para valores grandes de λ la región factible F^λ podría llegar a ser vacío.

Observemos en la Figura 2.1 el efecto de λ sobre el nivel de incerteza del conjunto robusto.

³Ver Definición 1.6 en página 11.

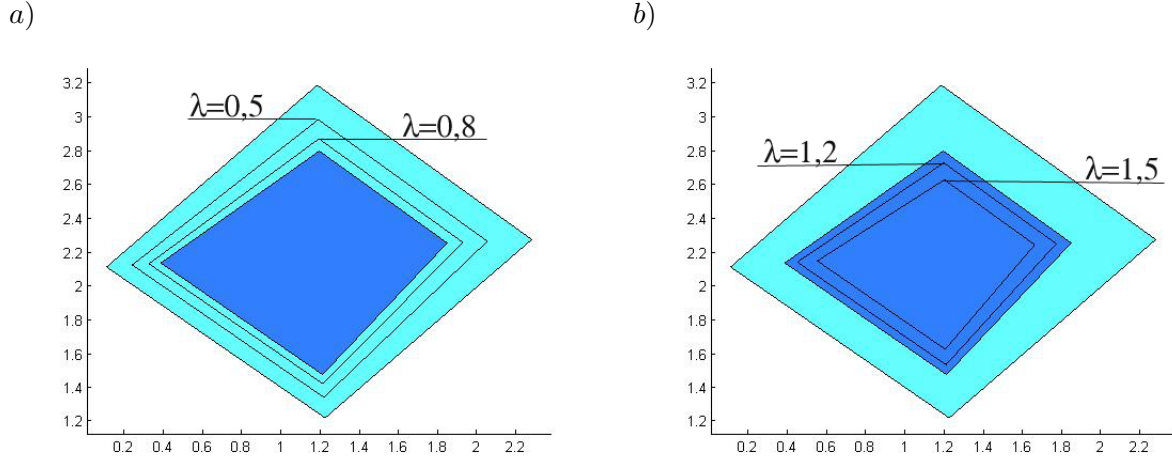


Figura 2.1: Efecto de λ como ponderador del nivel de incerteza. **a)** Disminución de la incerteza ($\lambda = 0,5$ y $\lambda = 0,8$). **b)** Aumento de la incerteza ($\lambda = 1,2$ y $\lambda = 1,5$).

Consideremos $x \in \{x \in \mathbb{R}^n / Ex = f, l \leq x \leq u\}$ y estudiemos las siguientes desigualdades

$$\varepsilon \left(\sum_{ij \in I} \delta_{ij} A_{ij} \right) x \leq b - Ax \quad \forall \delta \in \mathcal{U} \quad (2.1)$$

$$\lambda \varepsilon \left(\sum_{ij \in I} \delta_{ij} A_{ij} \right) x \leq b - Ax \quad \forall \delta \in \mathcal{U} \quad (2.2)$$

Satisfacer 2.1 ó 2.2 define la pertenencia a \mathcal{F} ó F^λ respectivamente. Tomemos $x \in F^0$ entonces $b - Ax \geq 0$ pero a priori no conocemos el signo de la incerteza (*lado izquierdo de 2.1*). Existen tres casos que debemos estudiar: x al interior, en la frontera y al exterior de \mathcal{F} . Recordemos que esta es una desigualdad vectorial, por lo que en su estudio debemos cuidar de observar todas sus componentes $i = 1, \dots, m_1$.

- Si $x \notin \mathcal{F}$ entonces existe algún δ que no satisface 2.1. Aquellas componentes que violan la desigualdad tendrán incerteza positiva, y entonces con algún $0 \leq \lambda < 1$ es posible satisfacer 2.2. Como $\lambda < 1$ las componentes que satisfacían 2.1 lo siguen haciendo (independiente del signo de la incerteza).

Luego, si variamos δ en \mathcal{U} el *mejor* λ nos dice *cuán lejos estaba x de ser robusto factible*.

- Si $x \in \text{int}(\mathcal{F})$ entonces 2.1 se satisface con holgura. Buscaremos llevar alguna componente a ser activa. En este caso la incerteza puede ser negativa o positiva. Se tendrá que
 - algún $\lambda > 1$ permite satisfacer 2.2 si la incerteza es positiva, y
 - algún $-\lambda$ permite satisfacer 2.2 si la incerteza es negativa. Si además ésta es lo suficientemente pequeña, entonces se podrá tener $\lambda > 1$

Así, si δ se mueve en \mathcal{U} el mejor λ nos dice *cuán robusto es x* .

- Si $x \in \partial\mathcal{F}$ entonces 2.1 se satisface con alguna componente activa y con $\lambda = 1$ se satisface 2.2. En este caso podemos decir que x es *débilmente robusto*, pues una pequeña variación adicional de los datos lo puede sacar de \mathcal{F} .

DEFINICIÓN 2.3. Para cada $x \in F^0$ se define el **margen de factibilidad** de x como

$$\lambda(x) := \sup \{ \lambda > 0 / (\forall \delta \in \mathcal{U}) A^\varepsilon(\pm\lambda\delta)x \leq b \}$$

Notemos que el signo \pm que acompaña a λ en la definición pierde importancia cuando el conjunto de incerteza es simétrico respecto al origen (de \mathbb{R}^I).

LEMA 2.4. Consideremos un conjunto de incerteza simétrico respecto al origen. Se tiene que

$$\lambda(x) = \sup \{ \lambda \geq 0 / (\forall \delta \in \mathcal{U}) A^\varepsilon(\lambda\delta)x \leq b \}$$

Demostración. En efecto, sea $\lambda \geq 0$ que satisface

$$\left| \begin{array}{ll} Ax + \lambda\varepsilon(\sum_{ij \in I} \delta_{ij} A_{ij})x \leq b & \forall \delta \in \mathcal{U} \\ Ax - \lambda\varepsilon(\sum_{ij \in I} \delta_{ij} A_{ij})x \leq b & \forall \delta \in \mathcal{U} \end{array} \right.$$

luego

$$\left| \begin{array}{ll} Ax + \lambda\varepsilon(\sum_{ij \in I} \delta_{ij} A_{ij})x \leq b & \forall \delta \in \mathcal{U} \\ Ax + \lambda\varepsilon(\sum_{ij \in I} -\delta_{ij} A_{ij})x \leq b & \forall \delta \in \mathcal{U} \end{array} \right.$$

o bien

$$\left| \begin{array}{ll} Ax + \lambda\varepsilon(\sum_{ij \in I} \delta_{ij} A_{ij})x \leq b & \forall \delta \in \mathcal{U} \\ Ax + \lambda\varepsilon(\sum_{ij \in I} \delta_{ij} A_{ij})x \leq b & \forall \delta \in -\mathcal{U} \end{array} \right.$$

Como \mathcal{U} es simétrico respecto al origen, es decir, $-\mathcal{U} = \mathcal{U}$ ambas restricciones coinciden, lo que permite concluir. ■

OBSERVACIÓN. Esta definición recoge lo discutido anteriormente. Podemos apreciar que $\lambda(x) > 1$ únicamente para aquellos puntos que son robusto factibles y que resisten una mayor incerteza. De la misma forma, un margen de factibilidad menor que uno está asociado a aquellos puntos que requieren una menor incerteza en los datos para poder ser robusto factible. La igualdad se tiene para los puntos que son robusto factibles, pero una pequeña variación del nivel de incerteza los puede sacar de \mathcal{F} . Como también se tendrá que

$$\lambda(x) = \sup \{ \lambda > 0 / (\forall \delta \in \mathcal{U}) A^{\pm\lambda\varepsilon}(\delta)x \leq b \}$$

Podemos entender $\lambda(x)\varepsilon$ como el máximo nivel de incerteza que permite a x ser robusto factible, es decir, $\lambda(\cdot)$ es una especie de termómetro de la confiabilidad, en otras palabras, establece una “distancia” entre x y $Fr(\mathcal{F})$.

Extendamos la definición de F^λ a

$$F^\lambda = \left\{ x \in \mathbb{R}^n / \begin{array}{l} A^\varepsilon(\pm\lambda\delta)x \leq b \quad \forall \delta \in \mathcal{U}, \\ Ex = e, \\ l \leq x \leq u \end{array} \right\}$$

Esta notación nos permite escribir

$$\lambda(x) = \sup\{\lambda \geq 0 / x \in F^\lambda\}$$

Vamos a generalizar el margen de factibilidad definiéndolo ahora para subconjuntos de F^0

DEFINICIÓN 2.5. Para cada $G \subseteq F^0$ definimos el margen de factibilidad de G como

$$\lambda(G) := \text{máx}\{\lambda \geq 0 / G \subseteq F^\lambda\}$$

Notemos que por definición se tiene que $x \in F^{\lambda(x)}$ y $G \subseteq F^{\lambda(G)}$. Entonces

$$\forall \lambda \leq \lambda(x) \quad x \in F^\lambda \quad \text{y} \quad \forall \lambda \leq \lambda(G) \quad G \subseteq F^\lambda$$

DEFINICIÓN 2.6. Sea G un subconjunto acotado de \mathbb{R}^n . Dados $0 \leq \rho \leq 1$ y $x \in G$. Llamaremos ρ -reducción de G en torno a x al conjunto

$$G_{\rho,x} := \{x\} + \rho(G - \{x\})$$

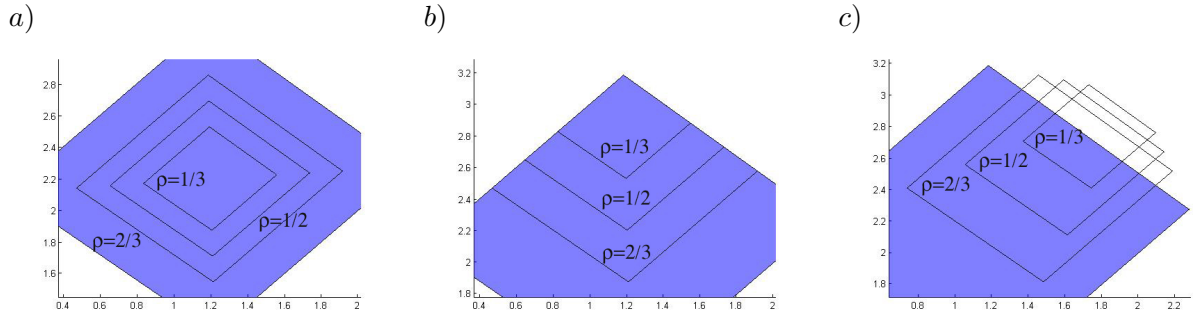


Figura 2.2: Reducciones de tamaño $\{\frac{2}{3}, \frac{1}{2}, \frac{1}{3}\}$ en torno a x . **a)** x al interior del poliedro. **b)** x un vértice del poliedro. **c)** x al exterior del poliedro, $(2, 3)$.

DEFINICIÓN 2.7. Dado $\rho \in (0, 1)$ diremos que el par (λ, ρ) es admisible si existe $\hat{x} \in F^0$ tal que $F_{\rho, \hat{x}}^0 \subseteq F^\lambda$. Denotaremos al conjunto de todos los pares admisibles (λ, ρ) por

$$\mathcal{A} := \left\{ (\lambda, \rho) / (\exists \hat{x} \in F^0) \quad F_{\rho, \hat{x}}^0 \subseteq F^\lambda \right\}$$

DEFINICIÓN 2.8. Sea x^0 óptimo nominal. Diremos que el par (λ, ρ) es débilmente admisible si existe $\hat{x} \in F^0$ tal que $\{\hat{x}, \hat{x} + \rho(x^0 - \hat{x})\} \subseteq F^\lambda$. Denotaremos al conjunto de todos los pares débilmente admisibles (λ, ρ) por

$$\mathcal{A}^* := \left\{ (\lambda, \rho) / (\exists \hat{x} \in F^0) \quad \{\hat{x}, \hat{x} + \rho(x^0 - \hat{x})\} \subseteq F^\lambda \right\}$$

LEMA 2.9. Sea $\rho \in (0, 1)$. Para todo $\hat{x} \in F^0$ existirá un par (λ, ρ) (débilmente) admisible.

Demostración. Sea $\hat{x} \in F^0$ se tiene que

$$\{\hat{x}, \hat{x} + \rho(x^0 - \hat{x})\} \subseteq F_{\rho, \hat{x}}^0 \subseteq F^0$$

La primera inclusión se tiene por definición de $F_{\rho, \hat{x}}^0$ y la segunda, puesto que $\hat{x} \in F^0$. De este modo

$$(0, \rho) \in \mathcal{A} \subseteq \mathcal{A}^*$$

■

OBSERVACIÓN. La admisibilidad de $\lambda > 1$ y $\rho \in (0, 1)$ dada por \hat{x} , asegura que en torno a \hat{x} existe una vecindad de puntos (dada por ρ) que soportan mayor incerteza que la necesaria para ser robusto factible, más aún, admiten un nivel de incerteza de $\lambda\varepsilon$. Veamos en la Figura 2.3 algunos casos de admisibilidad.

DEFINICIÓN 2.10. Dado el problema robusto

$$(\mathcal{P}_\varepsilon^*) \begin{cases} \text{mín} & c^T x \\ \text{s.a.} & x \in \mathcal{F} \end{cases}$$

definiremos el margen de factibilidad y el margen de factibilidad débil de $(\mathcal{P}_\varepsilon^*)$ como

$$\kappa_{\mathcal{U}} := \sup_{(\lambda, \rho) \in \mathcal{A}} \frac{\lambda\rho + 1}{1 - \rho} \quad , \quad \kappa_{\mathcal{U}}^* := \sup_{(\lambda, \rho) \in \mathcal{A}^*} \frac{\lambda\rho + 1}{1 - \rho}$$

OBSERVACIÓN. Notemos que $\frac{\lambda\rho+1}{1-\rho}$ tiene ambas derivadas parciales positivas. Además, para permitir ambas admisibilidades vemos que mientras más grande sea λ más pequeño debe ser ρ , y viceversa. Por ende, el supremo no es trivial.

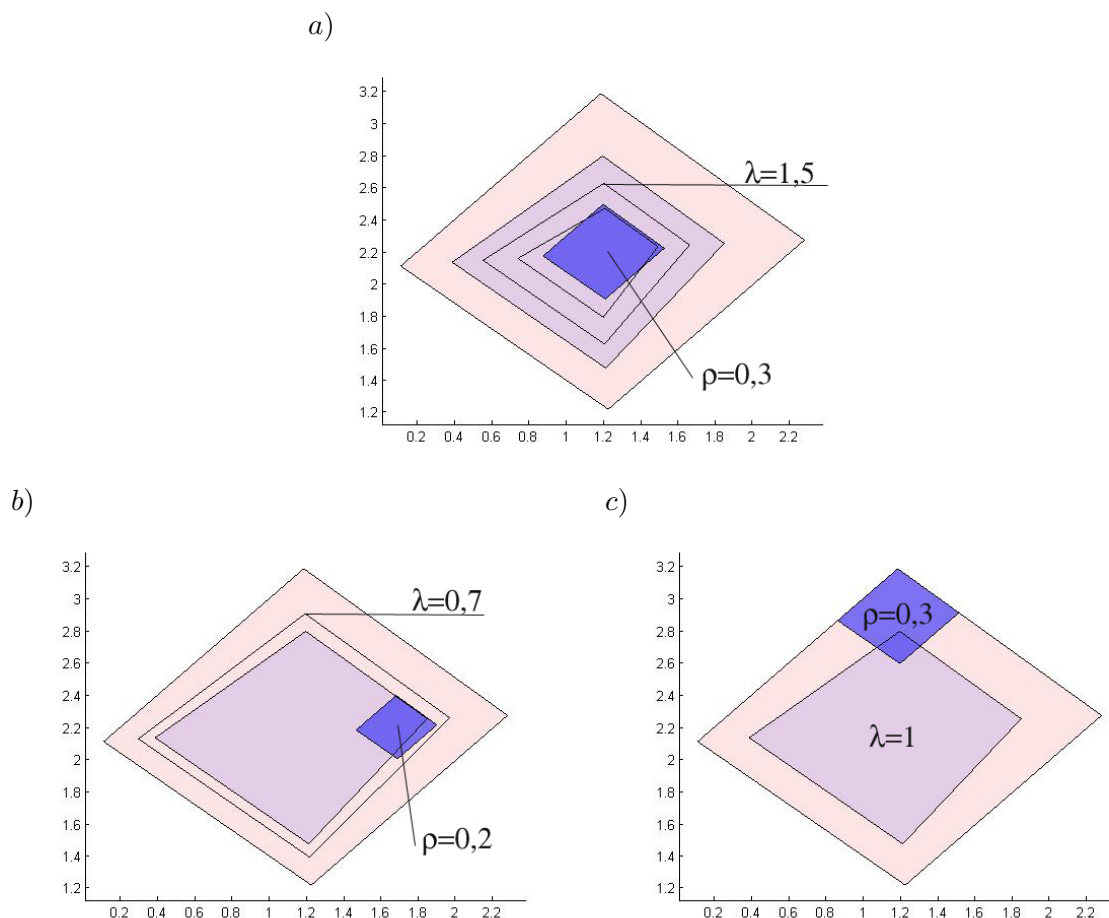


Figura 2.3: **a)** \hat{x} al centro admite λ grandes para ρ pequeño. Se observa que $(1,5 ; 0,3) \in \mathcal{A}$ y que \hat{x} no hace admisible a $(2 ; 0,3)$. **b)** Cuando la ρ -reducción sale de \mathcal{F} es necesario λ pequeño. Se observa que $(0,7 ; 0,2) \in \mathcal{A}$ y que \hat{x} no hace admisible a $(1 ; 0,2)$. **c)** Cuando \hat{x} es un vértice, solo $\lambda = 0$ permite la inclusión.

2.3. Modelo para la Estimación de la Sensibilidad del Valor Óptimo de un Problema Lineal con Incertidumbre Poliedral de sus Datos

A partir de las definiciones anteriores construiremos una estimación de la sensibilidad en términos de la función objetivo. Revisaremos a continuación la construcción teórica de esta estimación y describiremos el modelo matemático que permitirá calcularla.

2.3.1. Construcción de la Estimación

Consideremos el siguiente lema

LEMA 2.11. Sea (λ, ρ) admisible con $\lambda > 1$. Sea $\hat{x} \in F^0$ quien permite la admisibilidad, es decir, $F_{\rho, \hat{x}}^0 \subseteq F^\lambda$ entonces

$$c^T x^* - c^T x^0 \leq (1 - \omega)(c^T \hat{x} - c^T x^0)$$

con

$$\omega = \frac{\rho(\lambda + 1)}{\lambda\rho + 2 - \rho}$$

Demostración. Previamente demostraremos que $F_{\rho, \hat{x}}^0 \subset F^\lambda \implies F_{w, \hat{x}}^0 \subset \mathcal{F}$.

1. Sabemos que $F_{\rho, \hat{x}}^0 \subset F^\lambda$. Sean $x \in F^0$ y $\delta \in \mathcal{U}$ entonces

$$\begin{aligned} & A^\varepsilon(\pm\lambda\delta)[\hat{x} + \rho(x - \hat{x})] \leq b \\ \implies & \left(\frac{1}{\rho(\lambda-1)}\right) A^\varepsilon(\pm\lambda\delta)[\hat{x} + \rho(x - \hat{x})] \leq \left(\frac{1}{\rho(\lambda-1)}\right) b \\ \implies & \left(\frac{1}{\rho(\lambda-1)}\right) A^\varepsilon(\pm\lambda\delta)\hat{x} + \left(\frac{1}{\lambda-1}\right) A^\varepsilon(\pm\lambda\delta)(x - \hat{x}) \leq \left(\frac{1}{\rho(\lambda-1)}\right) b \end{aligned}$$

pero $\frac{1}{\rho(\lambda-1)} = \frac{1}{(\lambda-1)\omega} + \frac{1-\rho}{\rho(\lambda+1)}$, luego

$$\left(\frac{1}{(\lambda-1)\omega}\right) [A^\varepsilon(\pm\lambda\delta)\hat{x} + \omega A^\varepsilon(\pm\lambda\delta)(x - \hat{x})] + \left(\frac{1-\rho}{\rho(\lambda+1)}\right) A^\varepsilon(\pm\lambda\delta)\hat{x} \leq \left(\frac{1}{\rho(\lambda-1)}\right) b$$

2. $\hat{x} \in F_{\rho, \hat{x}}^0 \implies \hat{x} \in F^\lambda$ luego

$$\begin{aligned} & A^\varepsilon(\pm\lambda\delta)[\hat{x}] \leq b \\ \implies & \left(\frac{1-\rho}{\rho(\lambda+1)}\right) A^\varepsilon(\pm\lambda\delta)\hat{x} \leq \left(\frac{1-\rho}{\rho(\lambda+1)}\right) b \end{aligned}$$

sumando con la expresión anterior (usando la versión con signo contrario) obtenemos

$$\left(\frac{1}{(\lambda-1)\omega}\right) A^\varepsilon(\pm\lambda\delta)[\hat{x} + \omega(x - \hat{x})] + \left(\frac{2(1-\rho)}{\rho(\lambda+1)}\right) A\hat{x} \leq \left(\frac{1}{\rho(\lambda-1)} + \frac{1-\rho}{\rho(\lambda+1)}\right) b$$

3. Tenemos $Ax \leq b$ luego $A(\hat{x} + x - \hat{x}) \leq b$. Sumando con la expresión anterior

$$\left(\frac{1}{(\lambda-1)\omega}\right) A^\varepsilon(\pm\lambda\delta)[\hat{x} + \omega(x - \hat{x})] + A \left[\left(\frac{2(1-\rho)}{\rho(\lambda+1)} + 1\right) \hat{x} + (x - \hat{x}) \right] \leq \left(\frac{2\lambda - \rho\lambda + \rho}{\rho(\lambda-1)(\lambda+1)} + 1\right) b$$

usando $\frac{2(1-\rho)}{\rho(\lambda+1)} + 1 = \frac{1}{w}$

$$\begin{aligned}
 \Rightarrow & \quad \frac{1}{\omega} \left\{ \left(\frac{1}{\lambda-1} \right) A^\varepsilon(\pm\lambda\delta)[\hat{x} + \omega(x - \hat{x})] + A[\hat{x} + \omega(x - \hat{x})] \right\} \leq \left(\frac{\rho\lambda^2 + 2\lambda - \rho\lambda}{\rho(\lambda-1)(\lambda+1)} \right) b \\
 \Rightarrow & \quad \frac{1}{\omega} \left\{ \left(\frac{\lambda}{\lambda-1} \right) A[\hat{x} + \omega(x - \hat{x})] \pm \left(\frac{1}{\lambda-1} \right) \varepsilon\lambda \sum_{ij \in I} \delta_{ij} A_{ij}[\hat{x} + \omega(x - \hat{x})] \right\} \leq \left(\frac{\lambda(\rho\lambda + 2 - \rho)}{\rho(\lambda-1)(\lambda+1)} \right) b \\
 \Rightarrow & \quad \left(\frac{\lambda}{(\lambda-1)\omega} \right) \left\{ A[\hat{x} + \omega(x - \hat{x})] \pm \varepsilon \sum_{ij \in I} \delta_{ij} A_{ij}[\hat{x} + \omega(x - \hat{x})] \right\} \leq \left(\frac{\lambda(\rho\lambda + 2 - \rho)}{\rho(\lambda-1)(\lambda+1)} \right) b \\
 \Rightarrow & \quad A^\varepsilon(\delta)[\hat{x} + \omega(x - \hat{x})] \leq \omega \left(\frac{\rho\lambda + 2 - \rho}{\rho(\lambda+1)} \right) b \\
 \Rightarrow & \quad A^\varepsilon(\delta)[\hat{x} + \omega(x - \hat{x})] \leq b
 \end{aligned}$$

de la arbitrariedad de $\delta \in \mathcal{U}$ tenemos que

$$A^\varepsilon(\delta)[\hat{x} + \omega(x - \hat{x})] \leq b \quad \forall \delta \in \mathcal{U} \quad (2.3)$$

4. También

$$E(\hat{x} + \omega(x - \hat{x})) = E(\hat{x}) + \omega(E(x) - E(\hat{x})) = e + \omega(e - e) = e \quad (2.4)$$

Además se tiene que

$$l \leq \hat{x} + \omega(x - \hat{x}) \leq u \quad (2.5)$$

Finalmente, de (2.3), (2.4) y (2.5) se obtiene que

$$F_{\omega, \hat{x}}^0 \subset \mathcal{F}$$

Lo que quiere decir que todos los puntos F^0 ω -reducidos en torno a \hat{x} son robusto factible, en particular

$$\hat{x} + \omega(x^0 - \hat{x}) \in \mathcal{F}$$

luego

$$\begin{aligned}
 c^T x^* & \leq c^T (\hat{x} + \omega(x^0 - \hat{x})) \\
 & = c^T \hat{x} + \omega c^T x^0 - \omega c^T \hat{x} \\
 & = (1 - \omega) c^T \hat{x} + (1 - (1 - \omega)) c^T x^0 \\
 & = (1 - \omega) (c^T \hat{x} - c^T x^0) + c^T x^0
 \end{aligned}$$

Finalmente

$$c^T x^* - c^T x^0 \leq (1 - \omega) (c^T \hat{x} - c^T x^0)$$

■

TEOREMA 2.12. [*Estimación de Ben-Tal y Nemirovski*]

Dado el problema robusto $(\mathcal{P}_\varepsilon^*)$ se tiene que

$$c^T x^* - c^T x^0 \leq \frac{2}{\kappa_{\mathcal{U}} + 1} \left[\max_{x \in F^0} c^T x - c^T x^0 \right]$$

Demostración. Usando el lema anterior, para cada (λ, ρ) admisible y $\hat{x} \in F^0$ que permite su admisibilidad con $\lambda > 1$ se tiene que

$$\begin{aligned} c^T x^* - c^T x^0 &\leq \left(\frac{2(1-\rho)}{\lambda\rho + 2 - \rho} \right) (c^T \hat{x} - c^T x^0) \\ \implies c^T x^* - c^T x^0 &\leq \left(\frac{2}{\frac{\lambda\rho+1}{(1-\rho)} + 1} \right) (c^T \hat{x} - c^T x^0) \end{aligned}$$

La estimación más ajustada posible se obtiene para el par (λ, ρ) que determina el margen de factibilidad de $(\mathcal{P}_\varepsilon^*)$. Como no es posible conocer a priori el \hat{x} asociado a ese par admisible, la mejor cota que podemos dar del término $c^T \hat{x} - c^T x^0$ es la varianza de F^0 . Así, tomando supremo sobre todos los pares admisibles en la última desigualdad obtenemos

$$c^T x^* - c^T x^0 \leq \frac{2}{\kappa_{\mathcal{U}} + 1} \left[\max_{x \in F^0} c^T x - c^T x^0 \right]$$

■

Vemos que la demostración del Lema 2.11 se sostiene en el hecho que si $F_{\rho, \hat{x}}^0 \subset F^\lambda$ entonces $F_{w, \hat{x}}^0 \subset \mathcal{F}$. El valor de esta inclusión subyace en que solo dos puntos de $F_{\rho, \hat{x}}^0$ están incluidos en F^λ y no todo el conjunto, lo que nos hace pensar que el Lema 2.11 puede relajar sus hipótesis para conseguir un resultado similar.

TEOREMA 2.13. Sea (λ, ρ) débil admisible con $\lambda > 1$. Sea $\hat{x} \in F^0$ quien permite la admisibilidad, es decir, $\{\hat{x}, \hat{x} + \rho(x^0 - \hat{x})\}$ está incluido en F^λ entonces

$$c^T x^* - c^T x^0 \leq (1 - \omega)(c^T \hat{x} - c^T x^0)$$

Demostración. Idéntica a la demostración del Lema 2.11. Revisemos las generalidades de este caso:

1. Sabemos que $\hat{x} + \rho(x^0 - \hat{x}) \in F^\lambda$. Entonces la desigualdad obtenida en 1. es equivalente a

$$\left(\frac{1}{(\lambda - 1)\omega} \right) [A^\varepsilon(\pm\lambda\delta)\hat{x} + \omega A^\varepsilon(\pm\lambda\delta)(x^0 - \hat{x})] + \left(\frac{1 - \rho}{\rho(\lambda + 1)} \right) A^\varepsilon(\pm\lambda\delta)\hat{x} \leq \left(\frac{1}{\rho(\lambda - 1)} \right) b$$

2. Tenemos también que $\hat{x} \in F^\lambda$ luego la desigualdad obtenida en 2. es equivalente a

$$\left(\frac{1}{(\lambda-1)\omega}\right) A^\varepsilon(\pm\lambda\delta)[\hat{x} + \omega(x^0 - \hat{x})] + \left(\frac{2(1-\rho)}{\rho(\lambda+1)}\right) A\hat{x} \leq \left(\frac{1}{\rho(\lambda-1)} + \frac{1-\rho}{\rho(\lambda+1)}\right) b$$

3. Se tiene $Ax^0 \leq b$ luego $A(\hat{x} + x^0 - \hat{x}) \leq b$. Desarrollando como en 3. obtenemos

$$A^\varepsilon(\delta)[\hat{x} + \omega(x^0 - \hat{x})] \leq b \quad \forall \delta \in \mathcal{U} \quad (2.6)$$

4. También

$$E(\hat{x} + \omega(x^0 - \hat{x})) = E(\hat{x}) + \omega(E(x^0) - E(\hat{x})) = e + \omega(e - e) = e \quad (2.7)$$

Además se tiene que

$$l \leq \hat{x} + \omega(x^0 - \hat{x}) \leq u \quad (2.8)$$

Finalmente, de (2.6), (2.7) y (2.8) se obtiene que $\hat{x} + \omega(x^0 - \hat{x}) \in \mathcal{F}$ de donde se concluye de manera análoga. ■

COROLARIO 2.14. *Dado el problema robusto $(\mathcal{P}_\varepsilon^*)$. Entonces*

$$c^T x^* - c^T x^0 \leq \frac{2}{\kappa_{\mathcal{U}}^* + 1} \left[\max_{x \in F^0} c^T x - c^T x^0 \right]$$

Demostración. Viene del Teorema 2.12, a partir del Teorema 2.13 cambiando la noción de admisibilidad por admisibilidad débil. ■

Con este último resultado entregamos la estimación que a continuación calcularemos. La complejidad de este modelo radica en que el supremo que define $\kappa_{\mathcal{U}}^*$ esta sobre un conjunto \mathcal{A}^* que no se conocemos a priori.

2.3.2. Modelamiento de la Estimación

El cálculo de la estimación requiere conocer el conjunto de pares débilmente admisibles. Para determinar si el par (λ, ρ) es débilmente admisible o no, es necesario establecer la existencia de \hat{x} tal que $\{\hat{x}, \hat{x} + \omega(x^0 - \hat{x})\}$ esté contenido en F^λ . Para determinar esto resolveremos un problema equivalente que es producto de los siguientes lemas.

LEMA 2.15. *Si $(\lambda, \rho) \in \mathcal{A}^* \implies \forall 0 < \mu \leq \lambda \quad (\mu, \rho) \in \mathcal{A}^*$.*

Demostración. Basta encontrar una ρ -reducción que este contenida en F^μ . Sabemos que si $\mu \leq \lambda$ entonces $F^\lambda \subseteq F^\mu$, así $\{\hat{x}, \hat{x} + \omega(x^0 - \hat{x})\} \subseteq F^\lambda \subseteq F^\mu$, luego $(\mu, \rho) \in \mathcal{A}^*$. ■

A partir de este lema y lo que se observa en la Figura 2.3 resulta natural pensar que para cada ρ existe un λ máximo. Este lema nos sugiere la existencia de una aplicación $\rho \mapsto \lambda_\rho$ que describe la frontera de \mathcal{A} . En un grafo ρ v/s λ esta función determina completamente \mathcal{A}^* como muestra la Figura 2.4.

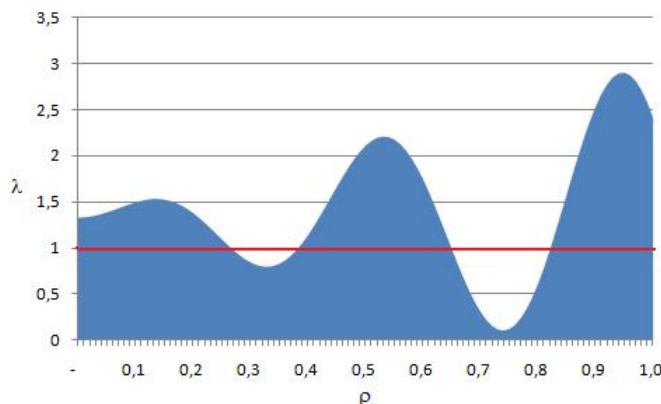


Figura 2.4: Esquema a priori de la región \mathcal{A}^* en un grafo ρ v/s λ . La línea $\lambda = 1$ separa aquellos pares que permiten la hipótesis del Teorema 2.13

LEMA 2.16. *Dado $\rho \in (0, 1)$. Entonces*

$$\lambda_\rho := \max_{x \in F^0} \lambda(\{x, x + \rho(x^0 - x)\})$$

es el mayor real tal que

$$\forall 0 < \mu \leq \lambda_\rho \quad (\mu, \rho) \in \mathcal{A}^*$$

Demostración. Debemos demostrar que

$$\mu^* := \max_{(\mu, \rho) \in \mathcal{A}^*} \mu = \lambda_\rho$$

Demostremos ambas desigualdades

1. Sabemos que

$$\forall x \in F^0 \quad (\lambda(\{x, x + \rho(x^0 - x)\}), \rho) \in \mathcal{A}^*$$

luego

$$\forall x \in F^0 \quad \lambda(\{x, x + \rho(x^0 - x)\}) \leq \mu^*$$

es decir $\lambda_\rho \leq \mu^*$.

2. Supongamos que $\lambda_\rho < \mu^*$ entonces $\exists \mu > \lambda_\rho$ tal que $(\mu, \rho) \in \mathcal{A}^*$, llamemos y a aquel vector que permite esta admisibilidad, es decir

$$\{y, y + \rho(x^0 - y)\} \subset F^\mu \tag{2.9}$$

pero $\lambda(\{y, y + \rho(x^0 - y)\}) \leq \lambda_\rho < \mu$ lo que es una contradicción, pues por definición $\lambda(\{y, y + \rho(x^0 - y)\})$ es el mayor real que permite la inclusión en 2.9, luego $\mu^* \leq \lambda_\rho$. ■

Notemos que si $\lambda_\rho \leq 1$ entonces no existen pares débilmente admisibles con segunda coordenada ρ que permitan el Teorema 2.13. De esto, obtenemos el siguiente resultado

TEOREMA 2.17. *Dado el problema robusto $(\mathcal{P}_\varepsilon^*)$. Su conjunto de pares admisibles que permiten el Teorema 2.13 esta dado por*

$$\mathcal{A} = \bigcup_{0 < \rho < 1} (1, \lambda_\rho] \times \{\rho\}$$

con λ_ρ el valor óptimo de

$$(\Lambda_\rho) \begin{cases} \text{máx} & \lambda \\ \text{s.a} & \\ & x \in F^\lambda \\ & x + \rho(x^0 - x) \in F^\lambda \end{cases}$$

(Convengamos que $(1, \lambda_\rho] = \phi$ si $\lambda_\rho \leq 1$).

Demostración. Es necesario demostrar ambas afirmaciones

1.

$$\begin{aligned} (\lambda, \rho) \in \mathcal{A}^* &\iff \exists \hat{x} \in F^0 \text{ tal que } \{\hat{x}, \hat{x} + \rho(x^0 - \hat{x})\} \subset F^\lambda \quad \wedge \quad \lambda > 1 \\ &\iff \exists \hat{x} \in F^0 \text{ tal que } \lambda(\{\hat{x}, \hat{x} + \rho(x^0 - \hat{x})\}) \geq \lambda > 1 \\ &\iff \text{máx}_{x \in F^0} \lambda(\{x, x + \rho(x^0 - x)\}) \geq \lambda > 1 \\ &\iff \lambda_\rho \geq \lambda > 1 \\ &\iff (\lambda, \rho) \in (1, \lambda_\rho] \times \{\rho\} \end{aligned}$$

2.

$$\begin{aligned} \text{máx}_{x \in F^0} \lambda(\{x, x + \rho(x^0 - x)\}) &= \text{máx}_{x \in F^0} \left\{ \text{máx} \lambda / \{x, x + \rho(x^0 - x)\} \subset F^\lambda \right\} \\ &\iff \begin{cases} \text{máx} & \lambda \\ \text{s.a} & \\ & \{x, x + \rho(x^0 - x)\} \subset F^\lambda \end{cases} \\ &\iff \begin{cases} \text{máx} & \lambda \\ \text{s.a} & \\ & x \in F^\lambda \\ & x + \rho(x^0 - x) \in F^\lambda \end{cases} \end{aligned}$$

■

De esta manera, para cada $\rho \in (0, 1)$ hemos construido un problema de optimización que encuentra

el par (λ, ρ) débilmente admisible con el mayor valor de λ posible. Como mencionamos anteriormente $\frac{\rho\lambda+1}{1-\rho}$ es creciente para ρ fijo, por lo que el máximo se alcanza justamente en (λ_ρ, ρ) de modo que

$$\kappa_{\mathcal{U}^*} = \sup_{\rho \in (0,1)} \frac{\rho\lambda_\rho + 1}{1 - \rho}$$

2.3.3. Cálculo de la Estimación

A continuación desarrollaremos el problema obtenido con el fin de obtener una versión que sea abordable computacionalmente.

$$\begin{aligned}
 (\Lambda_\rho) \quad & \left\{ \begin{array}{l} \text{máx} \quad \lambda \\ \text{s.a} \\ x \in F^\lambda \\ x + \rho(x^0 - x) \in F^\lambda \end{array} \right. \\
 \Leftrightarrow & \left\{ \begin{array}{l} \text{máx} \quad \lambda \\ \text{s.a} \\ A^\varepsilon(\pm\lambda\delta)x \leq b \quad \forall \delta \in \mathcal{U} \\ Ex = e \\ l \leq x \leq u \\ A^\varepsilon(\pm\lambda\delta)(x + \rho(x^0 - x)) \leq b \quad \forall \delta \in \mathcal{U} \\ E(x + \rho(x^0 - x)) = e \\ l \leq (x + \rho(x^0 - x)) \leq u \end{array} \right. \\
 \text{pero} & \left[\begin{array}{l} Ax^0 \leq b \\ Ex^0 = e \\ l \leq x^0 \leq u \end{array} \right. \\
 \Leftrightarrow & \left\{ \begin{array}{l} \text{máx} \quad \lambda \\ \text{s.a} \\ A^\varepsilon(\pm\lambda\delta)x \leq b \quad \forall \delta \in \mathcal{U} \\ A^\varepsilon(\pm\lambda\delta)(x + \rho(x^0 - x)) \leq b \quad \forall \delta \in \mathcal{U} \\ Ex = e \\ l \leq x \leq u \end{array} \right.
 \end{aligned}$$

$$\Leftrightarrow \left\{ \begin{array}{l} \text{máx} \\ \text{s.a} \\ \left[\begin{array}{l} A + \lambda\varepsilon \sum_{ij \in I} \delta_{ij} A_{ij} \\ A - \lambda\varepsilon \sum_{ij \in I} \delta_{ij} A_{ij} \end{array} \right] x \leq b \quad \forall \delta \in \mathcal{U} \\ \left[\begin{array}{l} A + \lambda\varepsilon \sum_{ij \in I} \delta_{ij} A_{ij} \\ A - \lambda\varepsilon \sum_{ij \in I} \delta_{ij} A_{ij} \end{array} \right] (x + \rho(x^0 - x)) \leq b \quad \forall \delta \in \mathcal{U} \\ Ex = e \\ l \leq x \leq u \end{array} \right.$$

Hasta ahora no hemos mencionado mayor detalle del conjunto de incerteza \mathcal{U} . A continuación detallaremos más sobre la descripción poliedral que haremos de \mathcal{U} y que permite incluir la alternativa de no tener todos los coeficientes inciertos.

SUPUESTO 2.18. *Para cada coeficiente a_{ij} de A existe un parámetro δ_{ij} que controla la incertidumbre de dicho coeficiente, incluso si el coeficiente es certero, hemos asumido que $\delta_{ij} = 0$. Podemos entender entonces que $\mathcal{U} \subseteq \mathcal{M}_{m_1 \times n}$, pero esto implica en muchos casos, la utilización innecesaria de recurso computacional valioso. Para efectos de implementación, es posible aprovechar la certeza de los coeficientes para disminuir la complejidad del algoritmo. La descripción poliedral que hemos tomado $\{M^i \delta_i \leq g^i\}$ asume que las restricciones inciertas son **independientes**. Si la i -ésima restricción no tiene coeficientes inciertos, entonces no es necesario definir M^i y g^i . Así mismo, para los coeficientes certeros de restricciones con incertidumbre, δ_{ij} es nulo, por lo que tampoco es necesario considerar aquellos δ_{ij} s.*

Denotaremos

$$\begin{array}{ll} \delta_i & \equiv \delta_i \quad \text{Vector columna} \\ I^1 & := \{i / \exists ij \in I\} \quad \text{Indices de restricciones con coeficientes inciertos} \\ I^2 & := \{1, \dots, m_1\} \setminus I^1 \quad \text{Indices de restricciones sin incertidumbre} \\ j(i) & := \{j / ij \in I\} \quad \text{Indices de variables de la restricción } i \text{ - esima} \\ & \quad \text{con coeficientes inciertos} \end{array}$$

Podemos escribir entonces

$$\delta_i \in \mathbb{R}^{|j(i)|} \quad \forall i \in I^1$$

y si se requieren $k(i)$ restricciones para describir poliedralmente a δ_i . entonces

$$M^i \in \mathcal{M}_{k(i) \times |j(i)|} \quad , \quad g^i \in \mathbb{R}^{k(i)} \quad \forall i \in I^1$$

OBSERVACIÓN. *En estricto rigor $\delta \in \mathcal{U}$ es una colección de $|I^1|$ vectores δ_i donde cada uno de*

ellos satisface $M^i \delta_i \leq g^i$.

Desarrollemos el primer bloque de restricciones

$$\begin{aligned} & \left[A + \lambda \varepsilon \sum_{ij \in I} \delta_{ij} A_{ij} \right] x \leq b \quad \forall \delta \in \mathcal{U} \\ \Leftrightarrow & \begin{cases} \left([A + \lambda \varepsilon \sum_{i \in I^1} \sum_{j \in j(i)} \delta_{ij} A_{ij}] x \right)_i \leq b_i & (\forall i \in I^1) (\forall \delta : M^i \delta_i \leq g^i) \\ A_i x \leq b_i & (\forall i \in I^2) \end{cases} \end{aligned}$$

De la primera desigualdad

$$\begin{aligned} \Leftrightarrow & [A_i + \lambda \varepsilon \sum_{j \in j(i)} \delta_{ij} (A_{ij})_i] x \leq b_i & (\forall i \in I^1) (\forall \delta : M^i \delta_i \leq g^i) \\ \Leftrightarrow & \sum_{j \in j(i)} \delta_{ij} (A_{ij})_i x \leq \frac{1}{\lambda \varepsilon} [b_i - A_i x] & (\forall i \in I^1) (\forall \delta : M^i \delta_i \leq g^i) \\ \Rightarrow & \max_{M^i \delta_i \leq g^i} \sum_{j \in j(i)} \delta_{ij} (A_{ij})_i x \leq \frac{1}{\lambda \varepsilon} [b_i - A_i x] & (\forall i \in I^1) \end{aligned}$$

Notemos que para $i \in I^1$

$$\sum_{j \in j(i)} \delta_{ij} (A_{ij})_i x = \sum_{j \in j(i)} \delta_{ij} a_{ij} x_j = ((a_{ij} x_j)_{j \in j(i)})^T \delta_i$$

Para cada $i \in I^1$ el dual del problema de maximización se escribe

$$\begin{cases} \text{mín} & (g^i)^T y^i \\ \text{s.a} & \\ & (M^i)^T y^i = (a_{ij} x_j)_{j \in j(i)} \\ & y^i \geq 0 \quad (\in \mathbb{R}^{k(i)}) \end{cases}$$

Por el teorema de dualidad fuerte, el valor de ambas funciones objetivo deben coincidir. Luego, podemos reemplazar el bloque de restricciones por uno equivalente dado por

$$\left| \begin{array}{ll} A_i x + \lambda \varepsilon (g^i)^T y^i \leq b_i & \forall i \in I^1 \\ (a_{ij} x_j)_{j \in j(i)} - (M^i)^T y^i = 0 & \forall i \in I^1 \\ A_i x \leq b & \forall i \in I^2 \\ y^i \geq 0 & \end{array} \right.$$

Para el segundo bloque de restricciones se obtiene de forma análoga

$$\left\{ \begin{array}{ll} A_i x + \lambda \varepsilon (g^i)^T z^i \leq b_i & \forall i \in I^1 \\ (a_{ij} x_j)_{j \in j(i)} + (M^i)^T z^i = 0 & \forall i \in I^1 \\ A_i x \leq b & \forall i \in I^2 \\ z^i \geq 0 & \end{array} \right.$$

Desarrollando de la misma forma el tercer y cuarto bloque se obtiene

$$(\Lambda_\rho) \Leftrightarrow \left\{ \begin{array}{ll} \text{máx} & \lambda \\ \text{s.a} & \\ & A_i x + \lambda \varepsilon (g^i)^T y^i \leq b_i \quad \forall i \in I^1 \\ & (a_{ij} x_j)_{j \in j(i)} - (M^i)^T y^i = 0 \quad \forall i \in I^1 \\ & A_i x + \lambda \varepsilon (g^i)^T z^i \leq b_i \quad \forall i \in I^1 \\ & (a_{ij} x_j)_{j \in j(i)} + (M^i)^T z^i = 0 \quad \forall i \in I^1 \\ & A_i (x + \rho(x^0 - x)) + \lambda \varepsilon (g^i)^T v^i \leq b_i \quad \forall i \in I^1 \\ & (a_{ij} (x_j + \rho(x_j^0 - x_j)))_{j \in j(i)} - (M^i)^T v^i = 0 \quad \forall i \in I^1 \\ & A_i (x + \rho(x^0 - x)) + \lambda \varepsilon (g^i)^T w^i \leq b_i \quad \forall i \in I^1 \\ & (a_{ij} (x_j + \rho(x_j^0 - x_j)))_{j \in j(i)} + (M^i)^T w^i = 0 \quad \forall i \in I^1 \\ & A_i x \leq b \quad \forall i \in I^2 \\ & E x = e \\ & l \leq x \leq u \\ & y^i, z^i, w^i, v^i \geq 0 \quad \forall i \in I^1 \end{array} \right.$$

TEOREMA 2.19. Si $\lambda_\rho > 0$ entonces

$$(\Lambda_\rho) \Leftrightarrow \left\{ \begin{array}{ll} \text{máx} & \lambda \\ \text{s.a} & \\ & A_i x + \lambda \varepsilon (g^i)^T y^i \leq b_i \quad \forall i \in I^1 \\ & (a_{ij} x_j)_{j \in j(i)} - (M^i)^T y^i = 0 \quad \forall i \in I^1 \\ & A_i x + \lambda \varepsilon (g^i)^T z^i \leq b_i \quad \forall i \in I^1 \\ & (a_{ij} x_j)_{j \in j(i)} + (M^i)^T z^i = 0 \quad \forall i \in I^1 \\ & A_i (x + \rho(x^0 - x)) + \lambda \varepsilon (g^i)^T v^i \leq b_i \quad \forall i \in I^1 \\ & (a_{ij} (x_j + \rho(x_j^0 - x_j)))_{j \in j(i)} - (M^i)^T v^i = 0 \quad \forall i \in I^1 \\ & A_i (x + \rho(x^0 - x)) + \lambda \varepsilon (g^i)^T w^i \leq b_i \quad \forall i \in I^1 \\ & (a_{ij} (x_j + \rho(x_j^0 - x_j)))_{j \in j(i)} + (M^i)^T w^i = 0 \quad \forall i \in I^1 \\ & A_i x \leq b \quad \forall i \in I^2 \\ & E x = e \\ & l \leq x \leq u \\ & y^i, z^i, w^i, v^i \geq 0 \quad \forall i \in I^1 \end{array} \right.$$

Demostración. Viene del desarrollo anterior. ■

COROLARIO 2.20. *Si \mathcal{U} es simétrico, entonces*

$$(\Lambda_\rho) \Leftrightarrow \left\{ \begin{array}{ll} \text{máx} & \lambda \\ \text{s. a} & \\ & A_i \cdot x + \lambda \varepsilon (g^i)^T y^i \leq b_i \quad \forall i \in I^1 \\ & (a_{ij} x_j)_{j \in j(i)} - (M^i)^T y^i = 0 \quad \forall i \in I^1 \\ & A_i \cdot (x + \rho(x^0 - x)) + \lambda \varepsilon (g^i)^T v^i \leq b_i \quad \forall i \in I^1 \\ & (a_{ij}(x_j + \rho(x_j^0 - x_j)))_{j \in j(i)} - (M^i)^T v^i = 0 \quad \forall i \in I^1 \\ & A_i x \leq b \quad \forall i \in I^2 \\ & Ex = e \\ & l \leq x \leq u \\ & y^i, v^i \geq 0 \quad \forall i \in I^1 \end{array} \right.$$

Demostración. Directo del Teorema 2.19 y el Lema 2.4. ■

COROLARIO 2.21. [Estimación de la Sensibilidad]

Dado el problema robusto $(\mathcal{P}_\varepsilon^)$. Sea (λ_ρ, x_ρ) solución de (Λ_ρ) . Entonces*

$$c^T x^* - c^T x^0 \leq \left(\frac{2(1-\rho)}{\lambda_\rho \rho + 2 - \rho} \right) [c^T x_\rho^* - c^T x^0]$$

o bien

$$c^T x^* - c^T x^0 \leq \frac{2}{\kappa_{\mathcal{U}}^* + 1} [c^T x_\rho^* - c^T x^0]$$

Demostración. Viene de la demostración del Teorema 2.12, tomando mínimo sobre $\rho \in (0, 1)$ en la desigualdad que entrega el Teorema 2.13. El Teorema 2.19 entrega aquél $\hat{x} = x_\rho^*$ que permite la admisibilidad débil de (λ_ρ, ρ) para el mejor ρ . ■

Este resultado nos permitirá construir un algoritmo (descrito en el Capítulo 4) que calcule estimaciones de la sensibilidad para la función objetivo de problemas lineales, posterior a la elección de un Modelo de Incerteza poliedral.

Capítulo 3

Dos Caracterizaciones de Incertidumbre Poliedral

Un parámetro poco confiable tiene consecuencias importantes en la calidad y confiabilidad de las soluciones de un problema de optimización. La Optimización Robusta es capaz, a pesar de la incerteza presente, de entregar soluciones confiables, pero antes debe conceptualizarse matemáticamente la variabilidad del parámetro.

3.1. Introducción

Con el fin de incorporar matemáticamente la incerteza de los datos en un modelo para que este sea capaz de encontrar soluciones robustas es necesario definir claramente los elementos de lo que hemos llamado Modelo de Incerteza. Este modelo busca conceptualizar teóricamente la incerteza en los datos con el fin de construir los criterios matemáticos que permiten modelar la factibilidad y la optimalidad robusta. El modelo de incerteza comprende cuatro elementos que se describen en el Capítulo 1, para poder afinar nuestro modelo, hacerlo comparable y finalmente utilizarlo en el grupo de problemas de control, profundizaremos estos elementos y algunos supuestos generales respecto a la incerteza presente en nuestros problemas.

Consideremos un problema lineal de la forma

$$(\mathcal{P}) \left\{ \begin{array}{ll} \text{mín} & c^T x \\ \text{s.a.} & \\ & Ax \leq b \\ & Ex = e \\ & l \leq x \leq u \end{array} \right.$$

con coeficientes inciertos en la matriz A . Ya hemos dicho que cualquier problema lineal incierto

puede ser reescrito como un problema de este tipo, por lo que este escenario es bastante general. Nuestro método de estimación de la sensibilidad requiere un conjunto de incertidumbre descrito poliedralmente. A partir de lo que discutiremos a continuación, esto también resulta ser en realidad una condición que se cumple en general. A continuación se enuncian los supuestos generales que son necesarios para la aplicación del método y que son parte de nuestro modelo de incerteza.

SUPUESTO 3.1. *La incerteza impactará sólo a la matriz de restricciones de desigualdad, y en ella, solamente a los coeficientes que sean inciertos, el modelador debe contar con un criterio que le permita decidir cuales coeficientes son inciertos y cuales no lo son. La incerteza afectará a todos los coeficientes inciertos de la misma forma, es decir, un único modelo de incerteza describe el impacto sobre todos los coeficientes a la vez*

SUPUESTO 3.2. *Como siempre debemos trabajar con información nominal de los problemas, es natural pensar que los posibles coeficientes serán una perturbación (porcentual) de estos, es decir*

$$\bar{a}_{ij} = (1 + \varepsilon\delta_{ij})a_{ij}$$

donde $\varepsilon > 0$ representa el nivel de incerteza de los datos.

En nuestro caso $|\delta_{ij}| < 1$ siempre, por lo que ε corresponde al porcentaje máximo que pueden llegar a variar los datos respecto a la versión nominal.

La descripción poliedral que requerimos para el parámetro δ nos permite asumir una gran cantidad de opciones respecto a la correlación que pudiesen tener pares o grupos de coeficientes. Nosotros consideraremos dos escenarios:

1. Todos los coeficientes inciertos tienen variaciones independientes entre sí.
2. Las restricciones son independientes entre sí, pero la incertidumbre conjunta en los coeficientes de una misma restricción está controlada por un parámetro, constante para todas las restricciones.

SUPUESTO 3.3. *No supondremos ningún criterio respecto al comportamiento de la variabilidad, salvo que el conjunto de incerteza debe ser un polítopo. Esto permite aplicar el modelo tanto a sistemas donde la incertidumbre tiene naturaleza estocástica, o donde la incerteza sólo parece estar acotada por un máximo y mínimo valor.*

Con la notación del capítulo anterior, recordemos que el problema (\mathcal{P}) con nivel de incertidumbre $\varepsilon > 0$ queda escrito como

$$(\mathcal{P}_\varepsilon) \left[\begin{array}{l} \text{mín} \quad c^T x \\ \text{s.a.} \\ A^\varepsilon(\delta)x \leq b \\ Ex = e \\ l \leq x \leq u \end{array} \right]_{\delta \in \mathcal{U}}$$

Anteriormente convenimos que usaríamos la siguiente estructura para describir el conjunto de incerteza

$$\mathcal{U} = \{\delta \in \mathcal{M}_{n \times m_1} / M^i \delta_i \leq g^i, i = 1, \dots, m_1\}$$

Luego, la contraparte robusta de \mathcal{P} está dada por

$$(\mathcal{P}^*) \left\{ \begin{array}{l} \text{mín} \\ \text{s.a.} \\ A^\varepsilon(\delta)x \leq b \\ M^i \delta_i \leq g^i \quad i = 1, \dots, m_1 \\ Ex = e \\ l \leq x \leq u \end{array} \right. \quad c^T x$$

En el capítulo anterior construimos para este problema un modelo que estima la sensibilidad de $c^T x$ frente a variaciones en los datos de A dados ε y \mathcal{U} (simétrico) como términos de la incertidumbre. Este modelo está dado por

$$c^T x^* - c^T x^0 \leq \frac{2}{\kappa_{\mathcal{U}}^* + 1} [c^T x_{\rho^*} - c^T x^0]$$

$$\kappa_{\mathcal{U}}^* = \sup_{\rho \in (0,1)} \frac{\rho \lambda_\rho + 1}{1 - \rho} = \frac{\rho^* \lambda_{\rho^*} + 1}{1 - \rho^*}$$

$$\lambda_\rho = \text{máx}_{\rho \in (0,1)} (\Lambda_\rho)$$

con (Λ_ρ) el valor óptimo de

$$\left\{ \begin{array}{l} \text{máx} \\ \text{s.a} \\ A_i x + \lambda \varepsilon (g^i)^T y^i \leq b_i \quad \forall i \in I^1 \\ (a_{ij} x_j)_{j \in j(i)} - (M^i)^T y^i = 0 \quad \forall i \in I^1 \\ A_i (x + \rho(x^0 - x)) + \lambda \varepsilon (g^i)^T v^i \leq b_i \quad \forall i \in I^1 \\ (a_{ij} (x_j + \rho(x_j^0 - x_j)))_{j \in j(i)} - (M^i)^T v^i = 0 \quad \forall i \in I^1 \\ A_i x \leq b \quad \forall i \in I^2 \\ Ex = e \\ l \leq x \leq u \\ y^i, v^i \geq 0 \quad \forall i \in I^1 \end{array} \right. \quad \lambda$$

Ver Definición 2.10 y Corolarios 2.14, 2.20 y 2.21. En lo que sigue aplicaremos este modelo a dos escenarios generales dados por bolas en \mathbb{R}^I con normas $\|\cdot\|_1$ y $\|\cdot\|_\infty$.

3.2. Dos Modelos de Sensibilidad con Incertidumbre Poliedral

La incertidumbre, es en cualquier ámbito, un elemento complejo y difícil de llevar a términos prácticos. Cuando se decide conceptualizar este elemento para trabajar sobre el efecto que este tiene en el sistema y sus respuestas, se recomienda como primer paso conocer el comportamiento que han manifestado históricamente los datos. Cuando esta información no existe, es necesario estimar en términos generales el nivel de incerteza que pudiese existir con el fin de elaborar un *rango* en el cual los datos pudiesen variar. Existen dos escenarios lo suficientemente generales que permiten aterrizar el concepto matemático de la incerteza, y que nos permitirán aplicar el modelo de estimación de sensibilidad a nuestro grupo de problemas de control.

Frente al desconocimiento total de la forma en que la incertidumbre afecta a los datos de un problema, la mejor estrategia es asumir que todos los coeficientes varían de forma independiente, y que ellos (los valores reales) son una variación porcentual de la versión conocida (o estimada) del parámetro. Este corresponde al primer escenario que desarrollaremos, y además es el caso considerado por Ben-Tal y Nemirovski en [4], estudio que será nuestro estándar de comparación.

Por otro lado, considerando que **una** restricción responde a **alguna** condición estructural de un sistema, es razonable considerar que la incertidumbre entre restricciones sea en ocasiones independiente. Más aun, es razonable pensar que es *difícil* que todos los coeficientes de una misma restricción varíen de la peor forma posible simultáneamente, es posible considerar que la incertidumbre sobre los coeficientes de una misma restricción está controlada por un nivel máximo. Este será nuestro segundo escenario.

Estos escenarios tienen una representación matemática que permite aplicarlos a una infinidad de casos. Estos corresponden a bolas de radio 1 entorno al origen de \mathbb{R}^I usando $\|\cdot\|_1$ y $\|\cdot\|_\infty$. Esto las hace de fácil representación lineal y son equivalentes, vía ponderación, a otros escenarios más complejos.

En la Figura 3.1 se puede ver como las dos descripciones del conjunto de incertidumbre afectan la geometría del problema original, lo que permite ver que un Modelo de Incerteza no es indiferente al conjunto \mathcal{U} considerado.

3.2.1. Modelo Bajo Incertidumbre Independiente en los Coeficientes

Al considerar que los posibles valores de los coeficientes inciertos son a lo más una variación del $\varepsilon\%$ del valor nominal, simplemente decimos que

$$\begin{aligned}\bar{a}_{ij} &= (1 + \varepsilon\delta_{ij})a_{ij} && \text{con} \\ -1 &\leq \delta_{ij} \leq 1 && \forall ij \in I\end{aligned}$$

La descripción de δ en este caso corresponde a la bola de radio 1 en torno al origen de \mathbb{R}^I considerando

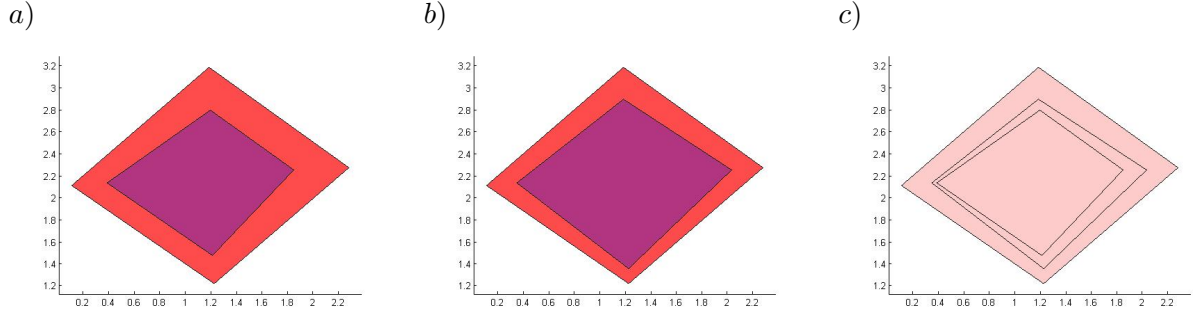


Figura 3.1: Poliedro $Ax \leq b$ versus \mathcal{F} . **a)** Incertidumbre de los coeficientes independiente. **b)** Incertidumbre de las restricciones independiente. **c)** Versus de los tres conjuntos.

la norma infinito. En la notación $\mathcal{U} = \{M^i \delta_i \leq g^i\}$ esto se obtiene con

$$M^i = \begin{bmatrix} I_{j(i)} \\ -I_{j(i)} \end{bmatrix}, \quad g = \begin{pmatrix} \mathbb{1}_{j(i)} \\ \mathbb{1}_{j(i)} \end{pmatrix}$$

Reemplazando en el modelo obtenemos

$$\left. \begin{array}{l} \text{máx} \\ \text{s.a} \end{array} \right\} \begin{array}{l} \lambda \\ \\ A_i \cdot x + \lambda \varepsilon \sum_{k=1}^{k(i)} y_k^i \leq b_i \quad \forall i \in I^1 \\ a_{ij} x_j - y_j^i + y_{j+j(i)}^i = 0 \quad \forall ij \in I \\ \\ A_i \cdot (x + \rho(x^0 - x)) + \lambda \varepsilon \sum_{k=1}^{k(i)} y_k^i v^i \leq b_i \quad \forall i \in I^1 \\ a_{ij} (x_j + \rho(x_j^0 - x_j)) - v_j^i + v_{j+j(i)}^i = 0 \quad \forall ij \in I \\ \\ A_i x \leq b \quad \forall i \in I^2 \\ \\ E x = e \\ \\ l \leq x \leq u \\ \\ y^i, v^i \geq 0 \quad \forall i \in I^1 \end{array}$$

Recordemos que $y^i \in \mathbb{R}^{k(i)}$, donde $k(i) = 2 \times j(i)$.

Este modelo es el mismo utilizado por Ben-Tal y Nemirovski para conceptualizar la incerteza de los

datos. Ellos calculan, bajo estos supuestos, el óptimo robusto (recordemos que ellos consideran una aproximación elipsoidal de \mathcal{F} , por lo que teóricamente el óptimo encontrado no es el óptimo robusto, sin embargo es una muy buena aproximación del valor óptimo) para los problemas *más sensibles* de la librería `NetLib` de modo que este será el modelo que implementaremos con el fin de evaluar la eficiencia de nuestro método. Estos resultados se describen en el Capítulo 4.

3.2.2. Modelo Bajo Incertidumbre Independiente en las Restricciones

El caso de restricciones con incertidumbre independiente entre sí, y cuya incertidumbre total por restricción (la suma de las incertidumbres de todos sus coeficientes) está acotada por 1, se describe como

$$\sum_j |\delta_{ij}| \leq 1 \quad \forall i = 1, \dots, m_1$$

o bien podemos escribir

$$\sum_j \gamma_j^i \leq 1 \quad , \quad -\gamma_j^i \leq \delta_{ij} \leq \gamma_j^i$$

Esto corresponde a la bola de radio 1 en en torno al origen de \mathbb{R}^I considerando la norma 1. En la notación $\mathcal{U} = \{M^i \delta_i \leq g^i\}$ se esto obtiene con

$$M^i = \begin{bmatrix} I_n & -I_n \\ -I_n & I_n \\ \vec{0}_n^T & \mathbf{1}_n^T \end{bmatrix} \quad , \quad g = \begin{pmatrix} \vec{0}_n \\ \vec{0}_n \\ 1 \end{pmatrix}$$

En este caso asumimos que $|I| = m_1 \times n$. Además $y^i \in \mathbb{R}^{2n}$. Reemplazando obtenemos

$$(\Lambda_\rho) \left\{ \begin{array}{ll} \text{máx} & \lambda \\ \text{s.a} & \\ & Ax \leq b \\ & A_i.(x + \rho(x^0 - x)) + \lambda\varepsilon(\sum_{k=1}^n \gamma_k^i y_k^i + \sum_{k=n}^{2n} \gamma_{k-n+1}^i y_k^i) \leq b_i \quad \forall i \in I^1 \\ & \sum_j \gamma_j^i \leq 1 \quad \forall i \in I^1 \\ & \sum_{j \in j(i)} (x + \rho(x^0 - x))_j (A_{ij})_i^T - (y_k - y_{k+n})_{k=1}^n = \vec{0} \quad \forall i \in I^1 \\ & Ex = e \\ & l \leq x \leq u \\ & y^i \geq 0 \quad \forall i \in I^1 \end{array} \right.$$

3.3. Comentarios para el Modelador

Conceptualizar la incerteza es una materia compleja y más aún, el modelo que se elija tendrá consecuencias en las respuestas del modelo. Un modelo con incertidumbre poliedral permite ser utilizado en una amplia gama de problemas y situaciones. No se ha encontrado en la literatura, otros modelos que permitan computar la sensibilidad o estabilidad del valor óptimo de problemas de optimización frente a variaciones de los datos, y este modelo parece ser la única herramienta disponible. En casos en que esta estimación pueda otorgar información clave al tomador de decisiones se recomienda la modelación a través de conjuntos de incertidumbre poliedrales.

Si el contexto del problema no lo permite, es posible construir otro modelo replicando la metodología que se utilizó en el Capítulo 2, pero perdiendo la linealidad del método. Ante cualquier circunstancia es necesario definir con antelación los cuatro elementos del Modelo de Incertidumbre, pues ellos, además de permitir la modelación, determinan fuertemente el tamaño del problema a resolver, la aplicabilidad computacional y la calidad de la estimación conseguida.

Capítulo 4

Estimación de la Sensibilidad de Problemas de Reales en NETLib

4.1. Introducción

Con el fin de validar nuestros resultados, pondremos en práctica la estimación teórica obtenida en el Corolario 2.21. La principal motivación de este capítulo es exponer las consideraciones que se tomaron para efectivamente computar la estimación en un grupo de problemas de control y las mejoras algorítmicas que se propusieron durante la misma experiencia.

Para implementar un algoritmo que estime la sensibilidad es necesario zanjar algunas consideraciones, cuya importancia ya hemos expuesto en capítulos anteriores. Estas consideraciones permiten además de la implementación, la comparación con resultados de otros estudios similares, y son:

- La elección de un grupo de problemas de control y la elección de una instancia nominal para cada uno de ellos. En rigor, si solo se cuenta con una instancia de cada problema, esos mismos datos pueden ser considerados como nominales.
- La definición de la certeza o incerteza de un dato. Consideraremos un dato incierto como aquel que, dada su cantidad de dígitos, *probablemente* no representa su valor fiel a la realidad.
- La descripción poliedral para describir la incerteza. Sabemos que la estimación depende directamente del modelo de incerteza elegido. En el Capítulo 3 se exponen dos descripciones generales que pueden ser aplicadas.

4.2. Aplicación de Nuestra Estimación

A continuación se describe en completitud el ejercicio de validación realizado: Se justifica la elección de los problemas estudiados y el modelo de incertidumbre. Se bosqueja el algoritmo utilizado. Y finalmente, se exponen los resultados obtenidos.

4.2.1. Librería de Problemas

Existen numerosas librerías de problemas lineales que han surgido a partir de casos reales de la ingeniería (NETLib, MALLab, NAGLib, IOptLib, COMPluib, etc). Para la realización de nuestra investigación, deseábamos contar con una librería de problemas que hubiese sido utilizada anteriormente en otros estudios relacionados con la misma materia. En [4] los autores evalúan todos los problemas de la librería NETLib con el fin de determinar cuáles eran los problemas *más sensibles* a la variación de sus datos. NETLib es un repositorio de problemas para uso científico y computacional mantenida por los Bell Laboratories de AT&T en la Universidad de Tennessee y el Oak Ridge National Laboratory. La estimación que construimos tiene la misma estructura que los problemas de esta librería:

$$(\mathcal{P}) \begin{cases} \text{mín} & c^T x \\ \text{s.a} & \\ & Ax \leq b \\ & Ex = e \\ & l \leq x \leq u \end{cases}$$

En [4] se estudia experimentalmente la infactibilidad que puede provocar la incerteza de los datos sobre los óptimos nominales, es decir, por cuánto puede llegar un óptimo a violar las restricciones de un problema que es sometido a variaciones de sus parámetros en la matriz A de la forma $a_{ij}(1 + \varepsilon\delta_{ij})$ donde $\varepsilon > 0$ representa el % de variación máxima y δ_{ij} un parámetro de control. Este modelo variacional se describe en la forma poliedral que nosotros hemos considerado.

En el estudio se consideró lo siguiente: Si bien el valor de un parámetro puede ser poco confiable, es muy poco probable que todos los coeficientes de una restricción sean falsos de forma simultánea, por lo que en el caso de existir violaciones a las restricciones del problema, estas no deben ser tan graves como lo serían en el peor caso. Se realizó una exploración estadística de las violaciones que puede tener un óptimo y de esto, se construyó un índice de *desconfiabilidad* que cuantifica cuan infactible puede llegar a ser $x \in \mathbb{R}^n$ producto de la incerteza (considerando variaciones del ε %). Este índice está dado por

$$\max_i \left(\frac{\max\{\text{Media}(\zeta_x^i) + \text{StDv}(\zeta_x^i); 0\}}{\max\{1; |b_i|\}} \times 100\% \right)$$

con

$$\zeta_x^i \equiv \max\{0, a_i^T x - b_i\}$$

y fue calculado para las soluciones óptimas de 90 problemas de la librería NETLib, tomando niveles de incerteza $\epsilon = 10^{-4}, 10^{-3}$ y 10^{-2} , y con \mathcal{U} descrito como se describe en la sección 3.2.1 (el caso de parámetros independientes). Si la desconfiabilidad de una solución tomaba valores mayores al 5% entonces esta era considerada *mala*. En el estudio, 27 de estos problemas presentaron malas soluciones nominales con nivel de incerteza $\epsilon = 10^{-2}$, de ellos, 19 mantuvieron su condición con $\epsilon = 10^{-4}$. En el Anexo B se muestran los resultados obtenidos para estos 27 problemas.

De estos 27 problemas, se seleccionaron como grupo de control para nuestros cálculos, aquellos problemas cuyo tiempo de ejecución en una máquina con procesador Intel COREi5 usando CPLEX 11.2 fuese menor a 60 minutos. Estos 18 problemas son: 25FV47, 80BAU3B, ADLITTLE, AFIRO, BNL2, BRANDY, CAPRI, CYCLE, E226, FINNIS, GREENBEA, KB2, MAROS, NESM, SCFXM1, SCFXM2, SCFXM3 y SHARE1B.

Por otro lado, la estimación propuesta depende del conjunto de coeficientes inciertos I , de modo que para realizar una comparación correcta es necesario siempre considerar el mismo conjunto I . Entenderemos que respecto a la certeza de los coeficientes, existen algunos casos directos, por ejemplo, los coeficientes con valores enteros probablemente son datos estructurales del problema y por ende son *confiables*. Para determinar los coeficientes inciertos de un problema el criterio utilizado por los autores, y el mismo que asumiremos nosotros, fue el siguiente:

DEFINICIÓN 4.1. *Los coeficientes inciertos de un problema lineal, serán aquellos coeficientes de las restricciones de desigualdad que no se pueden representar de la forma $\frac{p}{q}$ con $1 \leq q \leq 100$.*

En el Anexo B se muestran los tamaños de los problemas seleccionados y el número de coeficientes inciertos que se obtienen con este criterio.

4.2.2. Algoritmo

Necesitamos un procedimiento computacional que encuentre el *mejor* par débilmente admisible, es decir, que encuentre aquel $x_\rho \in F^{\lambda_\rho}$ que hace débilmente admisible al par (λ_ρ, ρ) el que a su vez, minimiza el valor de

$$\frac{2}{\left(\frac{\rho\lambda_\rho+1}{1-\rho}\right) + 1}$$

donde

$$\lambda_\rho = \max_{\rho \in (0,1)} (\Lambda_\rho)$$

con (Λ_ρ) descrito en 3.2.1. La estrategia natural para resolver este problema sería

- Discretizar el intervalo $(0, 1)$ en N puntos equidistantes ρ_i y resolver (Λ_{ρ_i}) para cada i .
- Encontrar el par $(\lambda_{\rho_i}, \rho_i)$ con $\lambda_{\rho_i} > 1$ que maximiza $\frac{\rho_i \lambda_{\rho_i} + 1}{1 - \rho_i}$.

Recordemos de la Figura 2.4 que a ciencia cierta no conocemos una relación entre λ_ρ y ρ que nos permita prever el comportamiento de $\kappa_{\mathcal{U}}^*$. Sin embargo, como se ve en la Figura 4.1, en todos los

problemas estudiados se observó que $\frac{\rho\lambda_\rho+1}{1-\rho}$ resultó ser decreciente¹ respecto a ρ . De esto, surgió una nueva idea: *¿Es necesario resolver los N problemas?* La Figura 4.1 sugiere la búsqueda de un ρ cerca de 1. Por otro lado, la Figura 4.2 muestra que λ_ρ como función de ρ es decreciente, de modo que los pares admisibles con primera coordenada mayor a 1, de existir, siempre se encontrarán al lado izquierdo del intervalo². Así, al recorrer el intervalo $(0, 1)$ hacia la izquierda el algoritmo sólo debe encontrar *el primer* par débilmente admisible (λ_ρ, ρ) con $\lambda_\rho \sim 1^+$.

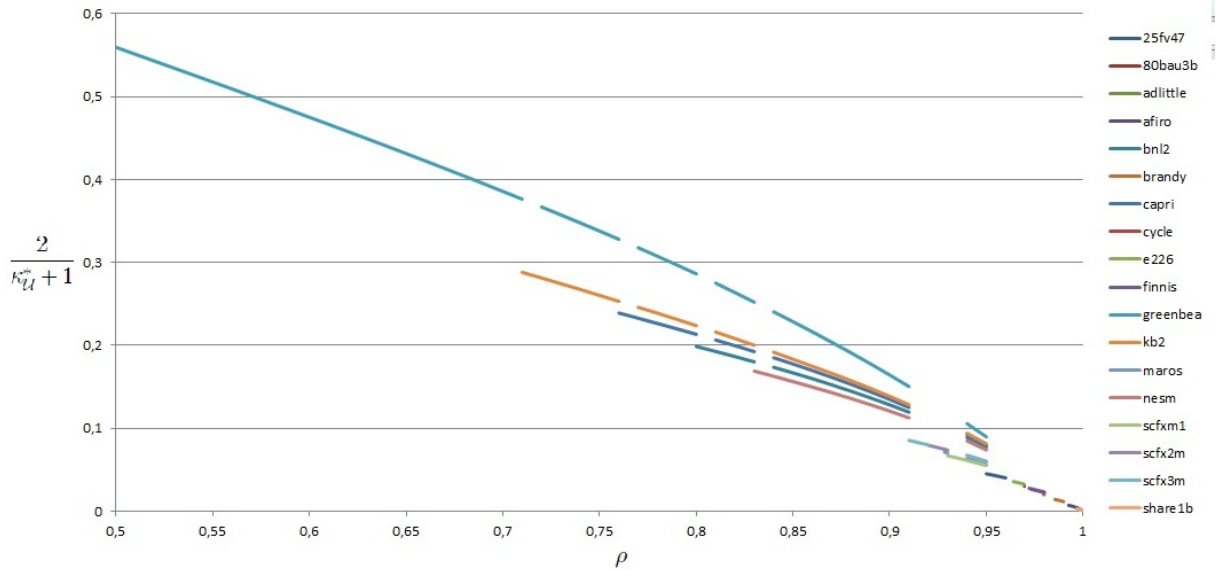


Figura 4.1: Gráfica del valor de $\frac{2}{k_i^* + 1}$ en función de ρ obtenido en la ejecución del algoritmo en el grupo de control

Además, se observó en los resultados que el ρ óptimo, en la gran mayoría de los casos, fue mayor a 0,6 por lo que para reducir la búsqueda, ésta debiese realizarse desde 1 hacia 0. Finalmente, con la intención de mejorar numéricamente la estimación, se aprovechó de refinar (tres veces) la discretización en las zonas del intervalo $(0, 1)$ donde el ρ óptimo estaría. En una primera instancia el intervalo se refinó en 100 puntos, luego buscó un par consecutivo de ρ_i 's de modo que tuvieran λ_ρ uno mayor y otro menor que 1. Esto indica que el ρ óptimo está entre ellos. Este procedimiento se repitió dos veces más, logrando precisión al 6^{to} decimal. El programa computacional quedó estructurado como se describe en el Anexo C.

4.2.3. Resultados Obtenidos

En las gráficas λ_ρ v/s ρ de las corridas del algoritmo para los problemas de control seleccionados incluidas la Figura 4.2, se puede observar cómo el algoritmo busca desde 1 hacia la izquierda y cómo

¹Esto no está demostrado

²Esto no está demostrado

refina una vez que ha dado con el *primer* $\lambda_\rho > 1$. Estas gráficas muestran la ejecución completa del algoritmo para tres problemas e incluye el detalle de la tercera refinación, en donde se puede observar cuando se alcanza el criterio de parada (En el Anexo D se muestran las 16 corridas). El efecto de escalera que se observa se debe a un error de truncación (en el 6^{to} decimal).

Recordemos algunas diferencias entre nuestra estimación propuesta y la estimación teórica señalada por Ben-Tal y Nemirovski. Ambas pretenden acotar la diferencia entre los valores optimos nominal y robusto. Cuando el nivel de incerteza sea ε notaremos el valor a estimar por

$$\Delta_\varepsilon = c^T x^* - c^T x^0$$

La dependencia de ε se encuentra implícita en el método de cálculo de $c^T x^*$. La estimación de Ben-Tal y Nemirovski plantea que dicha cantidad puede ser acotada superiormente por

$$\Delta_\varepsilon \leq \left(\frac{2}{\kappa_{\mathcal{U}} + 1} \right) \underbrace{\left(\max_{x \in F^0} c^T x - c^T x^0 \right)}_{:= VAR_{BN}}$$

en cambio, nuestra estimación³ propone la siguiente cota

$$\Delta_\varepsilon \leq \left(\frac{2}{\kappa_{\mathcal{U}}^* + 1} \right) \underbrace{\left(c^T x_\rho^* - c^T x^0 \right)}_{:= VAR}$$

En el Cuadro 4.1 se muestran los valores de VAR y VAR_{BN} para los 16 problemas de control. Ahí se observan algunas ventajas de la estimación propuesta: Por un lado nuestra cota permite estimar la sensibilidad de una mayor cantidad de problemas, y aún cuando la otra estimación es posible, nuestra estimación es más ajustada. En el Cuadro 4.2 se describen las estimaciones que se obtuvieron para tres niveles de incerteza y valor que se pretende estimar.

En el Cuadro 4.3 se muestran dos medidas diferentes que permiten cuantificar la calidad de la estimación calculada. Un primer indicador es

$$I_\varepsilon^1 := \frac{EST}{\max(1, \Delta)}$$

que señala cuantas veces la estimación *supera* al valor que deseamos estimar, cuando este es 0, el indicador vale simplemente el valor de la estimación. Un segundo indicador es

$$I_\varepsilon^2 := \frac{(EST - \Delta)}{|c^T x^0|}$$

que cuantifica el error cometido en relación al valor óptimo del problema nominal.

³Ver Corolario 2.21 en página 32

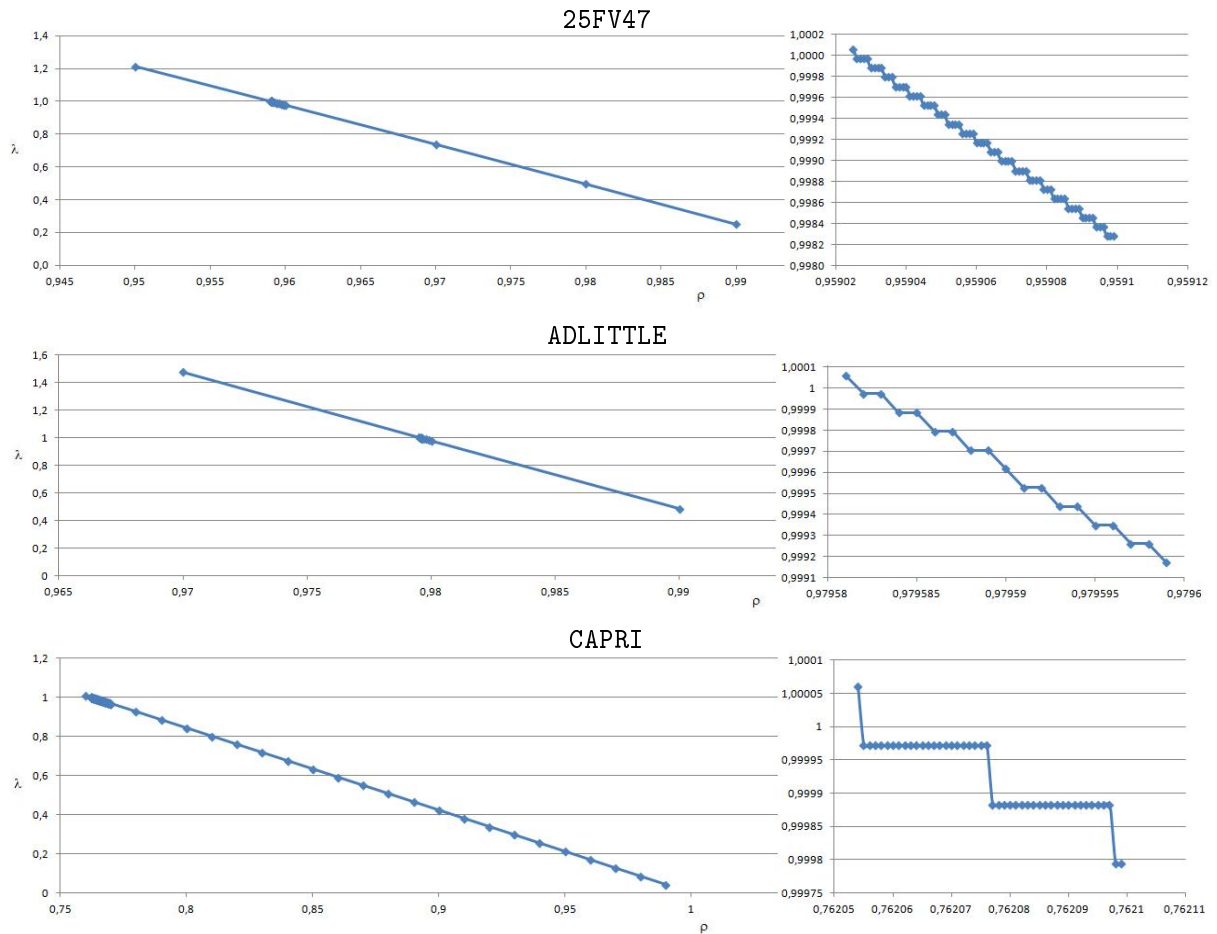


Figura 4.2: Gráficas λ_ρ v/s ρ que muestran el detalle de la ejecución del algoritmo con $\varepsilon = 0,01$ y la 3^{era} refinación.

Problema	VAR_{BN}	VAR
25FV47	∞	858,3440
8OBAU3B	∞	896.747,0000
ADLITTLE	∞	392.987,0000
AFIRO	3.903,04	464,7040
BNL2	∞	70,5629
BRANDY	∞	4.326,6200
CAPRI	∞	476,8020
CYCLE	1.001,04	172,0460
E226	104.556,75	20,4180
FINNIS	∞	362.247,0000
GREENBEA	∞	3.982.790,0000
KB2	1.749,90	338,1410
MAROS	47.440,30	72.373,4000
NESM	22.012.200,00	3.908.220,0000
SCFXM1	∞	2.133,7000
SCFXM2	∞	4.076,8300
SCFXM3	∞	6.101,0200
SHARE1B	151.151,80	901,4780

Cuadro 4.1: Dos tipos de varianzas calculadas para los problemas del grupo de control. VAR fue calculada considerando $\varepsilon = 0,01$

Problema	$\Delta_{0,0001}$	$EST_{0,0001}$	$\Delta_{0,001}$	$EST_{0,001}$	$\Delta_{0,01}$	$EST_{0,01}$
25FV47	0,2100	0,2933	2,110	3,411	21,10	35,17
80BAU3B	226,0000	271,9622	2.250,000	2.715,954	22.676,00	27.176,01
ADLITTLE	32,0000	80,4575	325,000	803,105	3.270,00	8.024,17
AFIRO	0,0060	0,0344	0,057	0,343	0,57	3,41
BNL2	0,0500	0,1456	0,570	1,455	5,75	13,92
BRANDY	0,0100	0,6226	0,050	6,172	0,52	62,43
CAPRI	1,1200	1,1246	11,180	11,255	112,80	113,45
CYCLE	0,0000	0,0481	0,000	0,551	0,00	4,59
E226	0,0052	0,0074	0,052	0,074	0,53	0,73
FINNIS	58,0000	99,7876	575,000	997,598	5.812,00	10.082,86
GREENBEA	37.200,0000	60.841,9967	372.600,000	550.169,484	3.876.600,00	3.902.254,00
KB2	0,0300	1,2213	0,270	12,147	3,29	96,21
MAROS	10,3000	18,6840	102,900	185,142	1.028,50	5.182,59
NESM	1.200,0000	10.974,0121	12.200,000	73.226,520	123.200,00	655.979,09
SCFXM1	0,5000	1,2542	5,600	12,599	62,40	135,92
SCFXM2	0,9000	2,9092	9,800	29,431	119,70	324,55
SCFXM3	1,4000	4,4471	13,900	44,986	177,10	490,32
SHARE1B	0,0000	0,0281	0,000	0,281	0,00	2,78

Cuadro 4.2: Valores estimados y a estimar de los problemas de seleccionados, considerando tres niveles de incerteza.

Problema	$I_{0,0001}^1$	$I_{0,001}^1$	$I_{0,01}^1$	$I_{0,0001}^2$	$I_{0,001}^2$	$I_{0,01}^2$
25FV47	1,397	1,617	1,667	0,00002	0,00024	0,00256
80BAU3B	1,203	1,207	1,198	0,00005	0,00047	0,00456
ADLITTLE	2,514	2,471	2,454	0,00021	0,00212	0,02108
AFIRO	5,725	6,017	5,993	0,00006	0,00062	0,00611
BNL2	2,912	2,552	2,420	0,00005	0,00049	0,00451
BRANDY	62,256	123,437	120,060	0,00040	0,00403	0,04077
CAPRI	1,004	1,007	1,006	0,00000	0,00003	0,00024
CYCLE	0,048	0,551	4,590	0,00920	0,10544	0,87828
E226	1,424	1,408	1,386	0,00012	0,00114	0,01091
FINNIS	1,720	1,735	1,735	0,00024	0,00245	0,02472
GREENBEA	1,636	1,477	1,007	0,00033	0,00245	0,00035
KB2	40,710	44,989	29,244	0,00068	0,00679	0,05310
MAROS	1,814	1,799	5,039	0,00014	0,00142	0,07154
NESM	9,145	6,002	5,325	0,00069	0,00434	0,03785
SCFXM1	2,508	2,250	2,178	0,00004	0,00038	0,00399
SCFXM2	3,232	3,003	2,711	0,00005	0,00054	0,00559
SCFXM3	3,176	3,236	2,769	0,00006	0,00057	0,00571
SHARE1B	0,028	0,281	2,783	0,00000	0,00000	0,00004

Cuadro 4.3: Indicadores de la calidad de la estimación calculada, para tres niveles de incerteza.

Capítulo 5

Análisis y Conclusiones

5.1. Introducción

Expuestos ya los resultados obtenidos en la ejecución del algoritmo implementado para el cómputo de la estimación de la sensibilidad, es momento de revisar las principales conclusiones que arrojó la experiencia. Haremos un análisis en detalle de los resultados, en comparación con estudios similares y propondremos algunas investigaciones adicionales que pueden realizarse en esta misma línea.

Para los análisis se consideraron dos estudios. En un primer lugar la investigación desarrollada por Ben-Tal y Nemirovski en [4], quienes en el 2000 estudiaron la sensibilidad de estos problemas, según se indica en la sección 4.2.1 y Anexo A. El segundo estudio considerado, fue el realizado por Ordoñez y Freund en [20], quienes en el 2004 calcularon el Número de Condicionamiento de estos problemas, medida que explica muchas de las características geométricas de los problemas lineales, en particular, la sensibilidad frente a variaciones de sus datos.

5.2. Análisis de los Resultados

Nuestra estimación está constituida por dos términos: el margen de factibilidad¹ κ_U^* y el tamaño del problema² VAR equivalentemente a como Ben-Tal y Nemirovski construyeron su cota teórica en función κ_U y VAR_{BN} como se describe en el Capítulo 2. En el Cuadro 4.1 se puede observar que esta estimación teórica no es aplicable en todos los casos, pues el término VAR_{BN} no siempre es acotado. Nuestra estimación repara esta desventaja, utilizando una cota superior más ajustada (VAR) en la demostración del Corolario 2.21. Por otro lado, cuando VAR_{BN} existe, en todos los problemas estudiados fue mayor que VAR lo que representa una posible ventaja de nuestra estimación.

Así como los términos de varianza contienen la información del tamaño o *diámetro* del problema,

¹Definido en página 21

²Definido en página 48

los términos dependientes del margen de factibilidad contienen la información de la sensibilidad estructural de la modelación (juntos estiman la sensibilidad del valor óptimo frente a variaciones de los datos). Este término es nuestro indicador de cuan sensible es en realidad la modelación y debiese ser explicado por otros indicadores afines a esta característica como el *Número de Condicionamiento* $C(d)$ o el *Índice de Desconfiabilidad* ID_ε (de las soluciones). Cuando sea calculado con nivel de incerteza ε a este término de la sensibilidad lo notaremos por

$$S_\varepsilon = \frac{2}{\kappa_{\mathcal{U}}^* + 1}$$

En el Cuadro 5.1 se muestran todos estos indicadores. Se realizaron estudios de correlación dos a dos de estas medidas, pero a pesar de lo esperado no existió correlación alguna. Esto se debe a que, si bien los indicadores miden de una u otra forma la *sensibilidad* del problema, estas observan características diferentes. El Número de Condicionamiento tiene una mirada global de la estructura geométrica del problema y contiene información de todas las restricciones, por otro lado, el Índice de Desconfiabilidad tiene información específica de la solución nominal, es posible que las restricciones inactivas no se vean afectadas por la incerteza cuando las activas sí, o viceversa, de modo que no guarda relación con indicadores globales. Por otro lado, la Sensibilidad Estructural examina el comportamiento de la región robusta en función del nivel de variabilidad de los datos, en forma global, por lo que contiene información distinta a la explicada por el Índice de Desconfiabilidad. Esto explicaría el hecho de que ningún par de variables estuvo relacionado en esta experiencia en particular.

Observemos ahora las estimaciones y el valor a estimar en el Cuadro 4.2, el error porcentual y el error relativo en el Cuadro 4.3 y el Número de Condicionamiento en el Cuadro 5.1. En general la estimación dio buenas cotas para la sensibilidad y salvo 3 excepciones obtuvo un error relativo menor al 8%. Esto quiere decir, que en relación a la magnitud del valor óptimo, la cota es buena. Se observa que a medida que el valor a estimar es mayor (problemas más sensibles) el Número de Condicionamiento crece, lamentablemente contamos con pocos datos de este indicador, pero suficientes para observar la tendencia.

Existen dos instancias con un comportamiento inusual: `CYCLE` y `SHARE1B`. A pesar que los datos fueron afectados por un 1% de variabilidad el valor óptimo robusto coincide con el valor óptimo nominal, es decir, el valor a estimar es 0. El algoritmo es capaz de identificar esto en el caso de `SHARE1B` pero no en el caso de `CYCLE`, el cual tiene muchos más coeficientes inciertos que `SHARE1B`³. Ambos problemas tienen Números de Condicionamiento altísimos, lo que indica que las restricciones inactivas se llevaron toda la carga de la incerteza.

Como la estimación presentaba buenos valores, quisimos observar cuan buen estimador fue en términos globales. Para ello calculamos la correlación entre el valor a estimar y el dato calculado por nuestro algoritmo. La Figura 5.1 presenta esta correlación en escala log – log, puesto que el valor a estimar, entre las diversas instancias, presentaba ordenes de magnitud bastante diferentes. El rees-

³Ver Anexo B

Problema	$C(d)$	$ID_{0,01}$	$S_{0,0001}$	$S_{0,001}$	$S_{0,01}$
25FV47	∞	1.620	0,000399	0,003998	0,040974
80BAU3B	8,5 - 8,7	8.420	0,000301	0,003004	0,030305
ADLITTLE	7,2	58	0,000207	0,002061	0,020418
AFIRO	3	50	0,000074	0,000738	0,007338
BNL2	8,9 - 9,0	34	0,001976	0,019749	0,197221
BRANDY	7,1 - 7,3	5	0,000146	0,001455	0,014430
CAPRI	7,9	390	0,002356	0,023580	0,237940
CYCLE	10,2	11.000	0,000265	0,002649	0,026680
E226	7,5 - 7,7	15	0,000360	0,003598	0,035978
FINNIS	∞	1.040	0,000282	0,002809	0,027834
GREENBEA	10,2	11.600	0,007598	0,077238	0,979779
KB2	7,7	2.680	0,002691	0,026776	0,284534
MAROS	∞	566	0,000582	0,005859	0,071609
NESM	-	20	0,001565	0,015704	0,167846
SCFXM1	9,3	9.460	0,000593	0,005963	0,063703
SCFXM2	9,6	9.460	0,000748	0,007516	0,079608
SCFXM3	9,8	9.460	0,000756	0,007593	0,080367
SHARE1B	11	25.700	0,000031	0,000309	0,003087

Cuadro 5.1: Número de Condicionamiento calculado por Ordoñez y Freund, como se señala en la Tabla 5 de [20] (pág. 20), Índice de Desconfiabilidad de las Soluciones Nominales calculado por Ben-Tal y Nemirovski como se señala en el Anexo B con nivel de incerteza $\varepsilon = 0,01$, y *Sensibilidad Estructural* calculada para tres niveles de incerteza, para todos los problemas de control.

calamiento permite observar una dependencia que podría parecer lineal, al menos, el valor a estimar explica el valor calculado con un $R^2 = 0,77$. Si efectivamente la dependencia fuese lineal en el logaritmo de los datos, entonces habría una dependencia potencial entre Δ y EST de la forma

$$EST = \alpha \Delta^\beta$$

5.2.1. Conclusiones

A través del desarrollo de este trabajo hemos explorado en detalle los principios fundamentales de la Optimización Robusta, hemos dado revista a diversas metodologías de resolución y hemos estudiado el impacto latente que tiene la incerteza de la información sobre la modelación matemática en problemas de decisión. Existe un sinnúmero de aplicaciones en las que metodologías conservadoras pueden o deben ser aplicadas, por lo que la teoría que hemos estudiado tiene un amplio campo de utilidad. La Optimización Robusta es una herramienta matemática contemporánea, que comenzó su verdadero desarrollo hace no más de 30 años, de modo que aún queda mucho espacio para lograr avances más profundos.

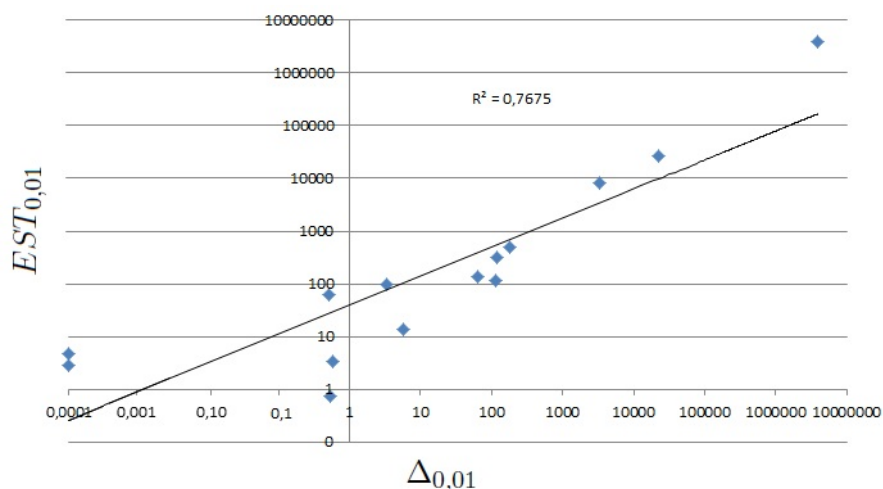


Figura 5.1: Correlación entre $\Delta_{0,01}$ y $EST_{0,01}$ en escala logarítmica. La línea de tendencia es sobre el logaritmo de los datos.

Si bien el objetivo del estudio fue desarrollar una medida computable, con el fin de poder elaborar análisis de sensibilidad a problemas reales, en el transcurso de la investigación apareció un elemento intrínseco, y transversal a toda esta teoría, que debió ser estudiado a cabalidad antes de permitirnos comenzar con la implementación del modelo estimador: el Modelo de Incertidumbre. La correcta *traducción* matemática de un ambiente de datos inciertos es vital para llevar a cabo cualquier tipo de cálculo o análisis. Solo después de estudiar y comprender las diversas formas en que la incerteza es modelada fuimos capaces de identificar los cuatro elementos comunes a todas las *tendencias* de modelación. Estos elementos⁴ dicen relación a como entendemos la incertidumbre y los análisis relacionados tendrán directa relación con ellos.

Durante la investigación quisimos ser lo más generales posibles, de modo de permitir la aplicación de nuestro algoritmo a un mayor número de problemas, de modo que elegimos un Modelo de Incerteza descrito a través de conjuntos poliedrales generalizados, permitiendo por ejemplo, la aplicación a dos tipos distintos de variación de los datos, lo que se detalló en el Capítulo 3. Nuestro aprendizaje y conclusión en este sentido fue que, producto de la existencia de diversas tendencias en la modelación de la incertidumbre, esta debe ser efectuada de forma transparente, explicitando todos los elementos que conforman el Modelo, esta es la única forma en que no se dé cabida a ningún tipo de abuso de lenguaje en la palabra “incierto”.

A partir de los resultados que arrojó nuestro algoritmo logramos apreciar dos aspectos de la teoría que son importantes al momento de enfrentarnos a una modelación de ambientes inciertos. Por un lado, las características geométricas de la región factible o de la región robusta, medidas por ejemplo a través del Número de Condicionamiento o el Margen de Factibilidad respectivamente, entregan sólo una idea *globalizada* de los eventuales efectos de la variación de los datos, la razón de esto es

⁴Detallados en el Capítulo 1

que sólo utilizan la información de las restricciones, y por ende no se enfocan específicamente en las zonas donde estará el óptimo, pudiendo esta tener una sensibilidad muy diferente al resto del conjunto factible. En definitiva, sabemos que la incerteza *ataca por todos lados* y que para medir el efecto de esta sobre el valor óptimo debiésemos contar con indicadores que consideren también la función objetivo para *localizar* las zonas de interés.

Por otro lado, los datos permiten concluir que la sensibilidad del valor óptimo tiene una dependencia lineal respecto al nivel de incerteza, al menos por tramos. Los tramos de esta dependencia se pueden explicar: Valores muy grandes de ε pueden hacer la región robusta infactible de modo que hay un límite superior para las variaciones posibles. Al *mover* ε entre 0 y este límite estamos considerando a la región factible nominal mientras sufre de modificaciones lineales en sus restricciones, es posible que algunas restricciones que eran redundantes dejen de serlo, incorporándose a la frontera de la nueva región y modificando la dependencia en el valor óptimo, generando un nuevo tramo de dependencia lineal. Aun así, en la mayoría de los casos estudiados, las restricciones redundantes no aparecieron. Una dependencia lineal del nivel de incerteza permitiría hacer análisis, sin invertir mayores recursos, en escenarios donde no se conoce la incerteza y se desea información amplia de las posibilidades.

Una desventaja (mejorable) de nuestro algoritmo es su tiempo de ejecución, la utilidad de la información que entrega permitiría ahorrar el tiempo que se consume en la ejecución de un algoritmo robusto, pero si dicha información tarda tanto o más tiempo en calcularse, la estimación no es practicable. Actualmente nuestro algoritmo es costoso en este sentido, sin embargo, al terminar entrega un punto inicial para algoritmos robustos que permitiría reducir bastante el tiempo de ejecución de posteriores cálculos, y además fue construido pensando en la exactitud de la estimación y no en el tiempo de ejecución, puesto que estábamos explorando esta teoría sin saber a priori si nuestra cota sería ajustada o no. De modo que el algoritmo tiene espacio para reducir el tiempo que consume.

Finalmente vemos que la herramienta que hemos construido calcula una buena cota superior de la sensibilidad. Las cotas calculadas fueron ajustadas en relación al orden de magnitud de los datos, lo que nos deja satisfechos en ese sentido, sabemos que aún hay mucho por mejorar en términos de eficiencia, pero logramos establecer al menos un algoritmo que permite medir la sensibilidad de los problemas y que entrega un buen punto inicial para la posterior ejecución de un algoritmo robusto.

5.3. Trabajo Futuro

A continuación se exponen las áreas en que la investigación puede ser mejorada o desarrollada en mayor profundidad.

5.3.1. Tiempo de Ejecución

La principal desventaja del algoritmo de estimación que hemos elaborado es su tiempo de ejecución, el objetivo inicial del trabajo desarrollado, además de conocer teóricamente una estimación de la

sensibilidad, es poder computar ese valor, y hacerlo de forma eficiente, con la idea de tener una herramienta que permita testear la practicidad de eventualmente ejecutar un algoritmo que entregue soluciones robustas. Sin embargo nuestra estimación, si bien entrega buenos valores, el tiempo que requiere el algoritmo implementado es demasiado alto. Es importante tener en cuenta que durante la construcción del algoritmo se privilegió la exactitud en el valor obtenido ante el tiempo de ejecución, de modo que aún queda oportunidad de mejorar en este sentido, en particular hay varias tareas concretas que hacer:

1. El algoritmo busca el primer ρ con $\lambda_\rho > 1$, comenzando desde 1 hacia la izquierda. En general esto resultó ser eficiente, pero hubieron instancias en que el ρ óptimo estuvo más cerca de 0 que de 1. Este problema se dio bastante en las segunda y tercera refinación, donde no existe criterio para evaluar cuál es el extremo más cercano del subintervalo. Si bien se recomienda realizar la búsqueda de izquierda a derecha, es recomendable también, cada cierto número de iteraciones, llevar una búsqueda en paralelo en sentido contrario.
2. En cada ρ el algoritmo busca λ_ρ (el λ más grande que es débilmente admisible), para ello hace una búsqueda exponencial resolviendo subproblemas de optimización cada vez más grandes, en el sentido que, el punto que hace admisible a (λ, ρ) también será punto factible del subproblema para evaluar el siguiente λ , de modo que cada solución puede ser utilizada como punto inicial para la resolución del subproblema siguiente, esto reducirá considerablemente la cantidad de tiempo de ejecución del algoritmo.
3. Del trabajo desarrollado en el Capítulo 2 y de la experiencia en la ejecución del algoritmo, es directo entender λ_ρ como una función de ρ . Si bien en un comienzo no se vislumbraban características de esta función, las figuras del Anexo D dan mayor claridad en este sentido. Existe una clara dependencia que incluso pareciese ser lineal - al menos una aproximación lineal sería bastante buena -. Por otro lado, el algoritmo se resume simplemente en la búsqueda de un ρ tal que su imagen por esta función sea mayor y muy cercana a 1. Todo esto sugiere la siguiente alternativa: Si conociésemos una buena estimación de esta dependencia, lineal o no, bastaría con obtener una preimagen de $1 + \epsilon$ por esta función para obtener el par débilmente admisible que buscamos.

La última, es la sugerencia más fuerte que se hace con el fin de obtener una segunda versión del algoritmo que calcule la estimación de forma eficiente, sin embargo la hipótesis de la dependencia es un argumento bastante fuerte que requiere demostración, o bien, un mayor estudio experimental para su validación.

5.3.2. Nivel de Incerteza y Restricciones Redundantes

Se puede observar en el Cuadro 4.3 que el error de la estimación en relación al valor óptimo del problema nominal I_ϵ^2 es proporcional al nivel de incerteza ϵ , de comprobarse esta idea, se presentaría

una ventaja significativa al momento de estudiar el impacto de la variación de los datos sobre problemas con nivel de incerteza desconocido. Sin embargo existe una consideración importante que es necesario tener en cuenta.

En este Cuadro se observa que existen algunos problemas (BRANDY, CYCLE, KB2, MAROS y NESM) que no se comportan como el resto, sus indicadores de calidad I_ε^1 e I_ε^2 o bien son notoriamente sin relación a los demás o al variar el nivel de incerteza no mantienen la misma tendencia. No existe una clara idea del porqué ocurre esto, pero tenemos una posible explicación de este comportamiento. Para la solución de un problema lineal, las restricciones pueden ser clasificadas como activas, inactivas y redundantes. En la mayoría de los problemas, las variaciones del 1% del valor de los datos modificarán las restricciones activas, literalmente *trasladando* la solución dentro de su vecindad, pero en ocasiones (lo que creemos que ocurrió en el caso de KB2, MAROS y NESM) las restricciones redundantes pueden llegar a ser activas, cambiando completamente la solución, y por ende modificando la estimación.

La dependencia de la estimación respecto a ε , o bien, la dependencia del valor óptimo $c^T x$ respecto a los coeficientes de A han sido una materia de amplio estudio, y el tema ha sido abordado desde varias áreas. Se sabe, y es consecuente con lo observado por nosotros, que esta dependencia es continua y lineal por tramos, donde los puntos de quiebre están dados por la activación de restricciones redundantes. Esta materia no deja de ser interesante y dejamos la puerta abierta para seguir investigando en esta dirección.

5.3.3. Solución Robusta

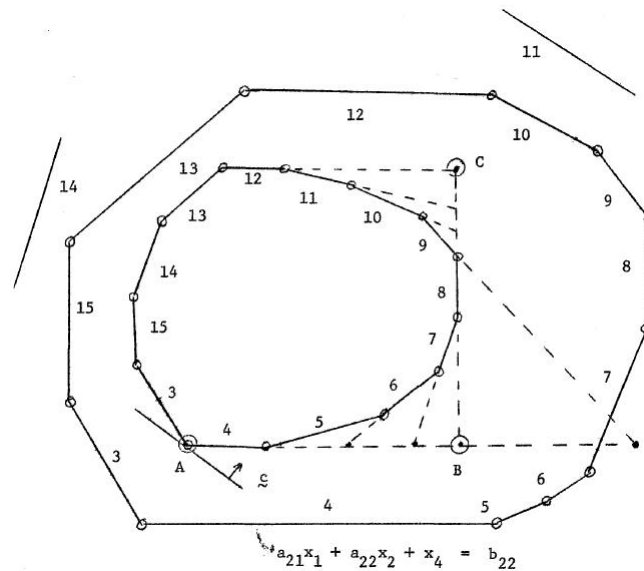
Como se ha señalado el algoritmo considera una secuencia de problemas que tienden al problema robusto. El criterio de parada permite al algoritmo entregar un par (λ_ρ, ρ) débilmente admisible dada la existencia de $x^\rho \in F^{\lambda_\rho}$ con $\lambda_\rho \sim 1+$. Recordemos que $\mathcal{F} = F^1$ de modo que al terminar, el algoritmo de estimación conoce un punto muy cercano a la región factible robusta, el cual es un buen punto de partida para un algoritmo que entregue soluciones robustas, mejorando su tiempo de ejecución.

De este modo, al construir un test que evalúa la conveniencia de un algoritmo robusto, al preferirlo, puede continuar iterando en la búsqueda del óptimo robusto aprovechando las iteraciones realizadas y ahorrando tiempo de ejecución. Esta herramienta no está implementada, pero el algoritmo estimador que se construyó es una subrutina de él bastante potente que puede ser aprovechada en este sentido.

Apéndice

Anexo A. Un primer bosquejo del impacto en la variación de los datos.

Esquema realizado por Gould de una región factible lineal con incertidumbre en los datos, en él se puede apreciar como restricciones redundantes (11 y 14) pueden dejar de serlo cuando se perturban los datos. Este esquema corresponde a la Figura 2 de [17] (pág. 18).



Anexo B. Grupo de Problemas de Control

Índice de Desconfiabilidad de Ben-Tal y Nemirovski calculado a los 27 problemas de la librería NETLib que resultaron tener las *peores* soluciones nominales al ser calculados para variaciones máximas del 1% y $\mathcal{U} = \{\delta : |\delta_{ij}| \leq 1 \quad \forall ij\}$. Este Cuadro resume la Tabla 2. de [4] (pág. 5) donde se señalan los índices para los tres niveles de incerteza.

Problema	$\epsilon = 0,01\%$	$\epsilon = 0,1\%$	$\epsilon = 1\%$
80BAU3B	84	842	8.420
25FV47	16	162	1.620
ADLITTLE	<5	6	58
AFIRO	<5	5	50
BNL2	<5	<5	34
BRANDY	<5	<5	5
CAPRI	<5	39	390
CYCLE	110	1.100	11.000
D2Q06C	1.150	11.500	115.000
E226	<5	<5	15
FFFFF800	<5	<5	8
FINNIS	10	104	1.040
GREENBEA	116	1.160	11.600
KB2	27	268	2.680
MAROS	6	57	566
NESM	<5	<5	20
PEROLD	34	339	3.390
PILOT	50	498	4.980
PILOT4	210.000	2.100.000	21.000.000
PILOT87	130	1.300	13.000
PILOTJA	46	463	4.630
PILOTNOV	69	694	6.940
PILOTWE	12.000	122.000	1.220.000
SCFXM1	95	946	9.460
SCFXM2	95	946	9.460
SCFXM3	95	946	9.460
SHARE1B	257	2.570	25.700

Número de filas, columnas y coeficientes inciertos (según la Definición 4.1) de los problemas seleccionados para el grupo de control del estudio.

Problema	# filas	# columnas	I
25FV47	822	1571	964
80BAU3B	2263	9799	5611
ADLITTLE	57	97	73
AFIRO	28	32	18
BNL2	2325	3489	3178
BRANDY	221	249	107
CAPRI	272	353	257
CYCLE	1904	2857	7054
E226	224	282	579
FINNIS	498	614	1189
GREENBEA	2393	5405	6096
KB2	44	41	107
MAROS	847	1443	866
NESM	663	2923	188
SCFXM1	331	457	402
SCFXM2	661	914	804
SCFXM3	991	1371	1206
SHARE1B	118	225	26

Anexo C. Algoritmo Implementado

Procedimiento de Cálculo de la Estimación

```

# Lambda>1(r) busca el par (l,x) tal que (l,r) es el par débilmente
# admisible por x, con segundo coordenada r y mayor primera
# coordenada (mayor que 1).

# Lambda(r,l) busca el punto x que hace a (l,r) débilmente
# admisible. Si no lo es, el problema es infactible.

# tres refinaciones para encontrar el primer rho con 3 decimales
let N=10
foreach r in (1,0)//N
    if Lambda>1(r).unfeasible then continue else {save r, break}
let N=10*N
foreach r in (r+0,1,r)//N
    if Lambda>1(r).unfeasible then continue else {save r, break}
let N=10*N
foreach r in (r+0,01,r)//N
    if Lambda>1(r).unfeasible then continue else {save r, break}

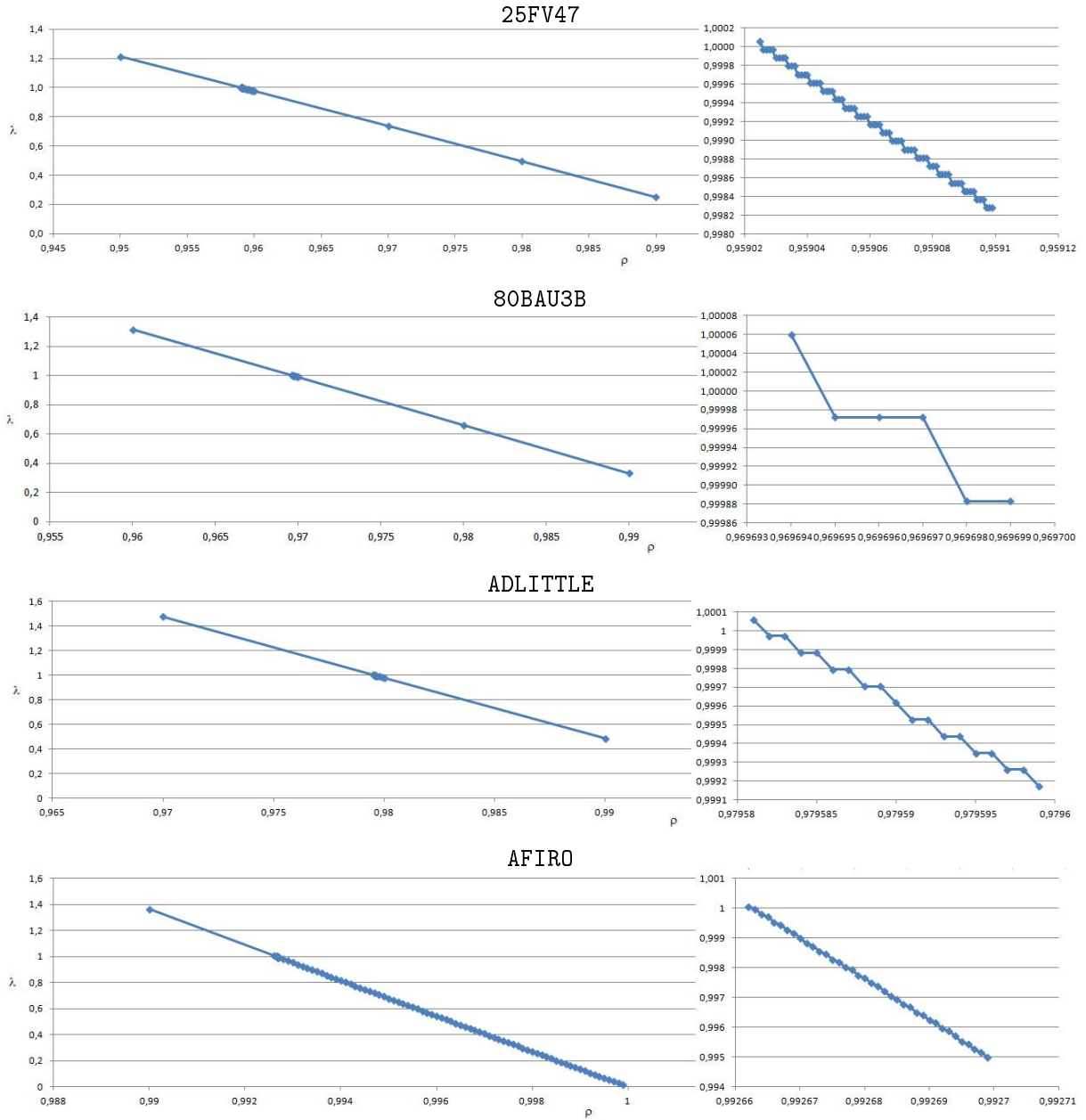
# Lambda>1(r) es un problema no lineal. Para resolverlo hacemos una
# búsqueda exponencial de l
let l=1,001
while Lambda(r,l).feasible
    l=2*l
let M=100
foreach l in (l/2,1)//M
    if Lambda(r,l).feasible then continue else {save l, break}
let L=l-(l-l/2)/M

return EST=2/(1+((Lr+1)/(1-r)))

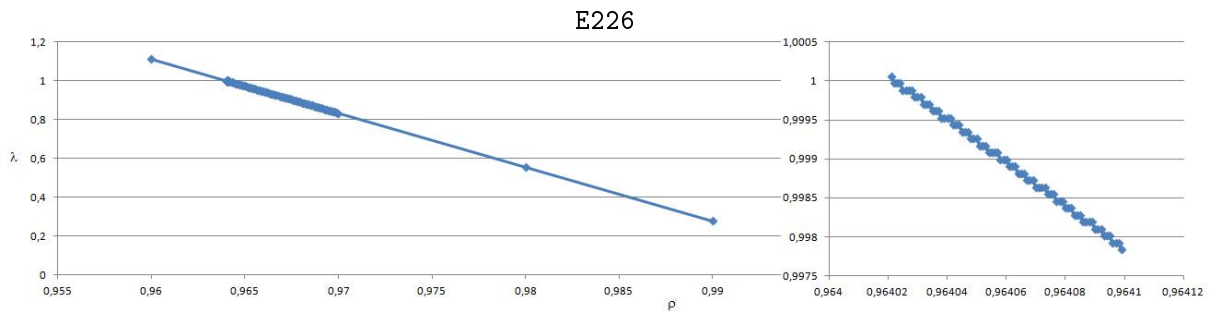
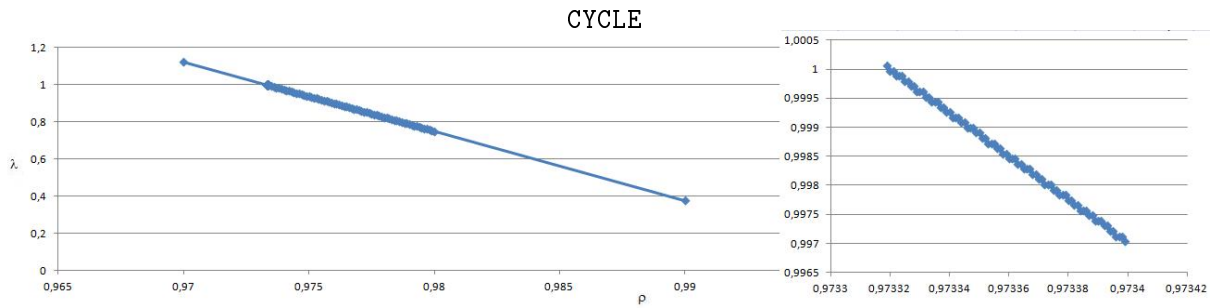
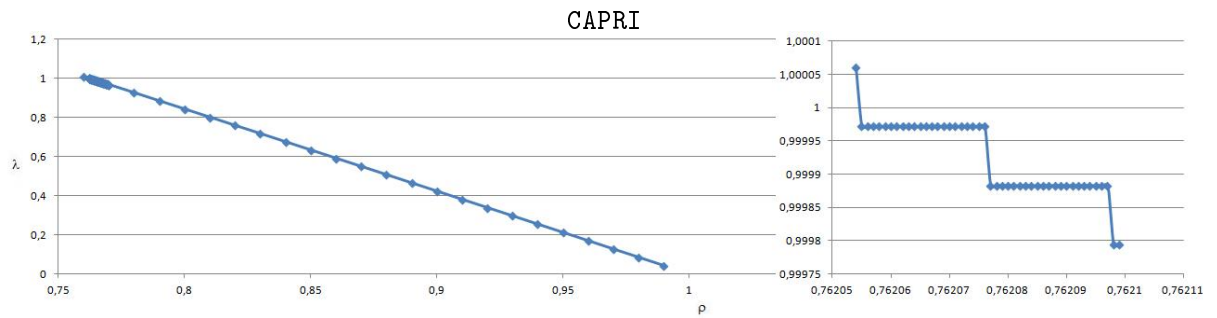
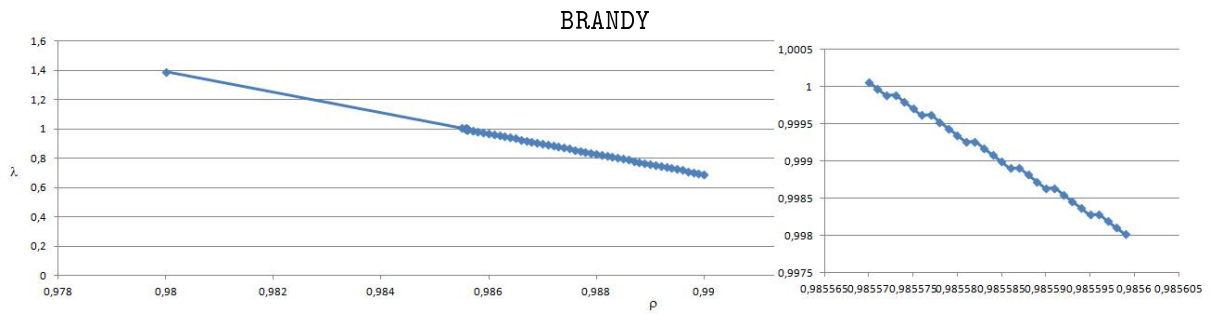
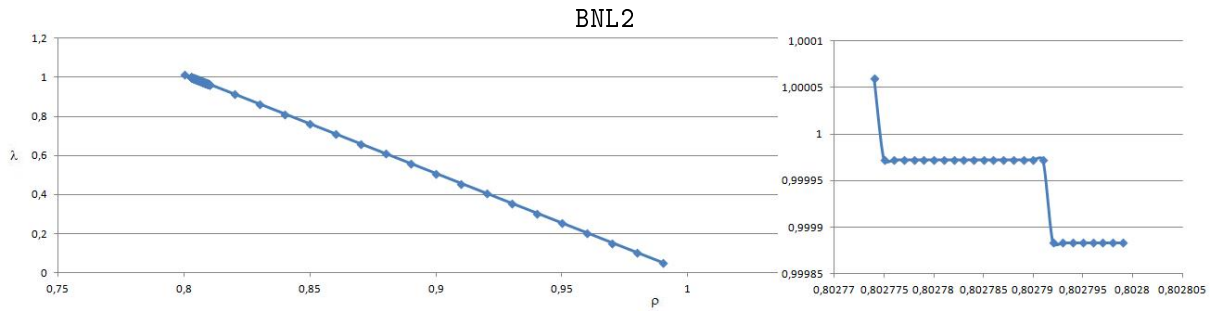
```

Anexo D. Aplicación del algoritmo en problemas de NETLib

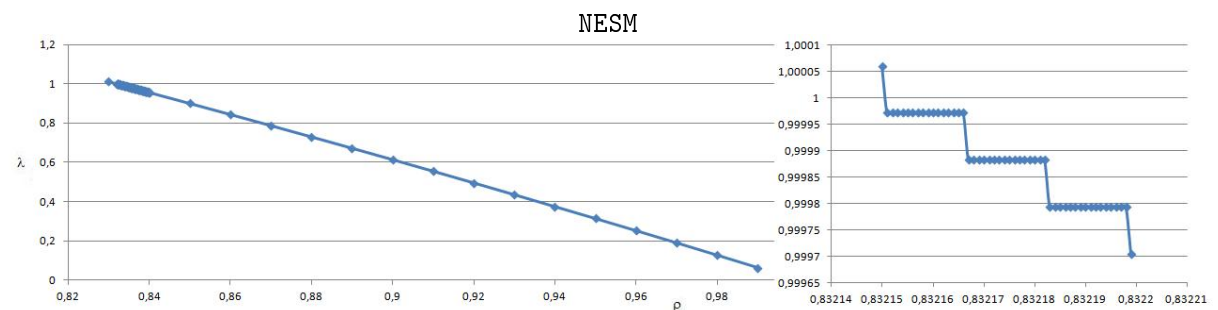
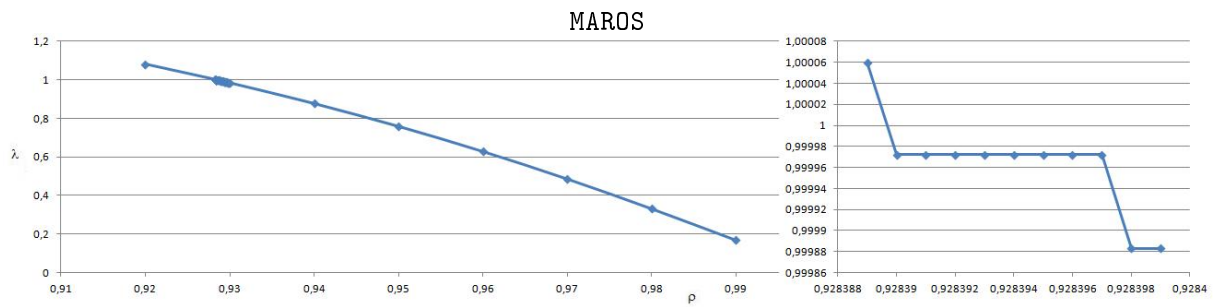
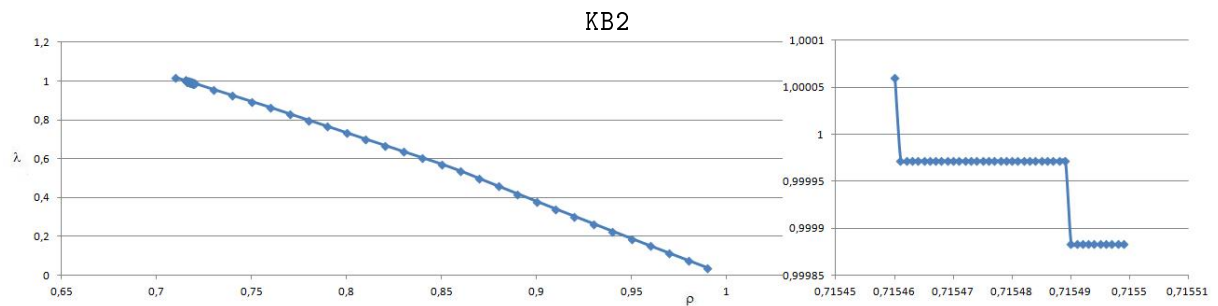
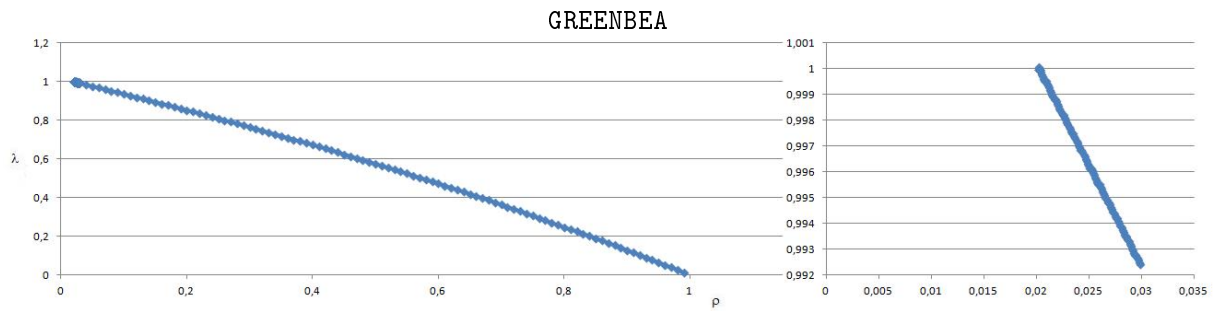
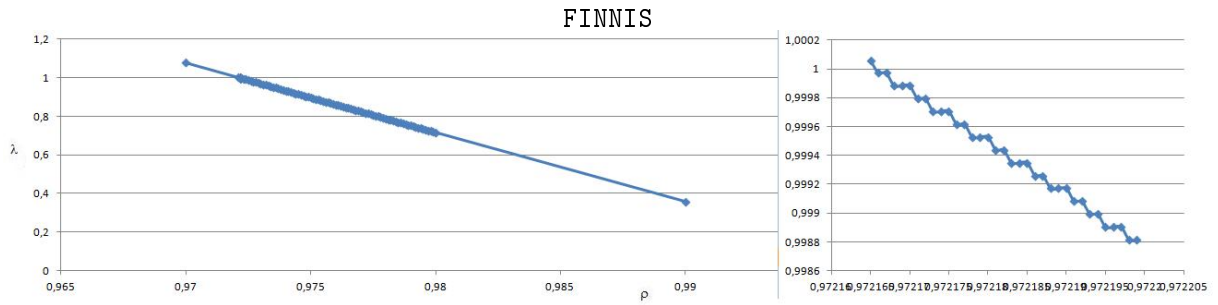
Gráficas λ_ρ v/s ρ que muestran los pares por los que pasó el algoritmo mientras se ejecutó considerando un nivel de incerteza $\varepsilon = 0,01$. Se muestra el detalle de la tercera refinación y el momento en que alcanza el criterio de parada, es decir, cuando λ_ρ supera el valor 1.



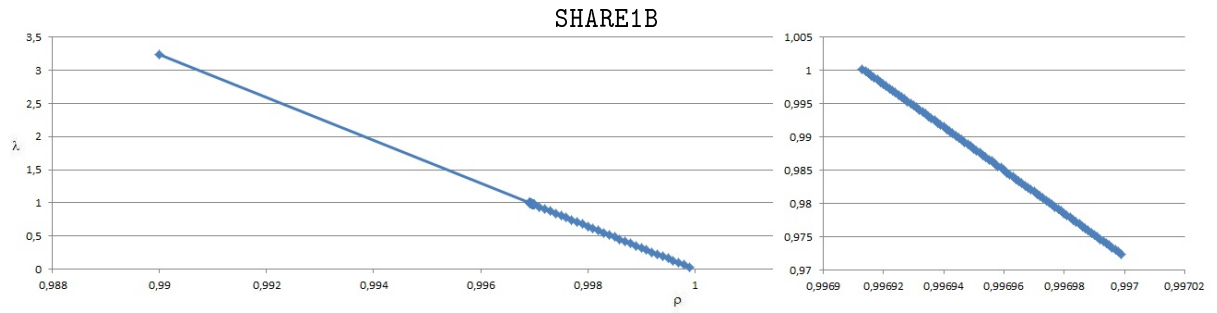
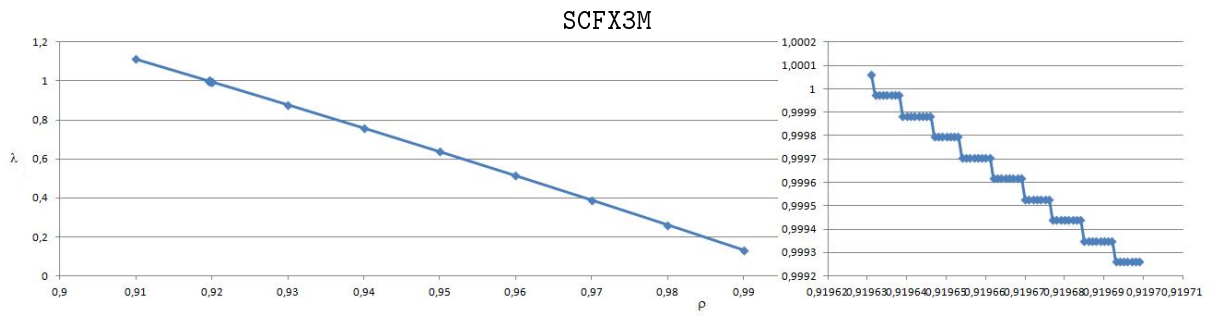
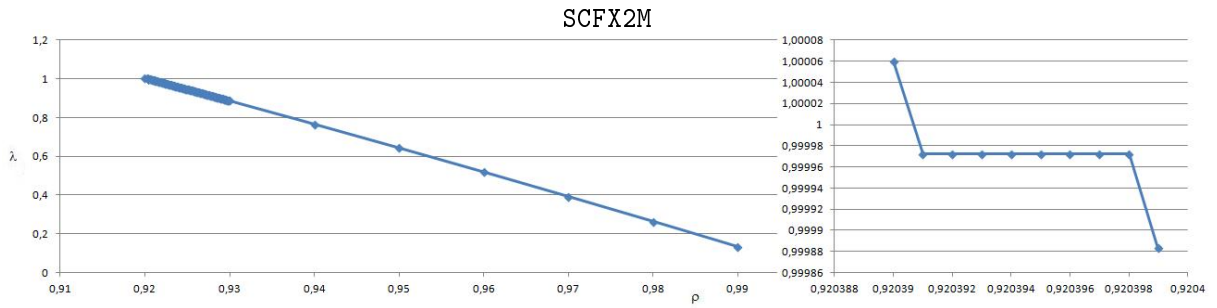
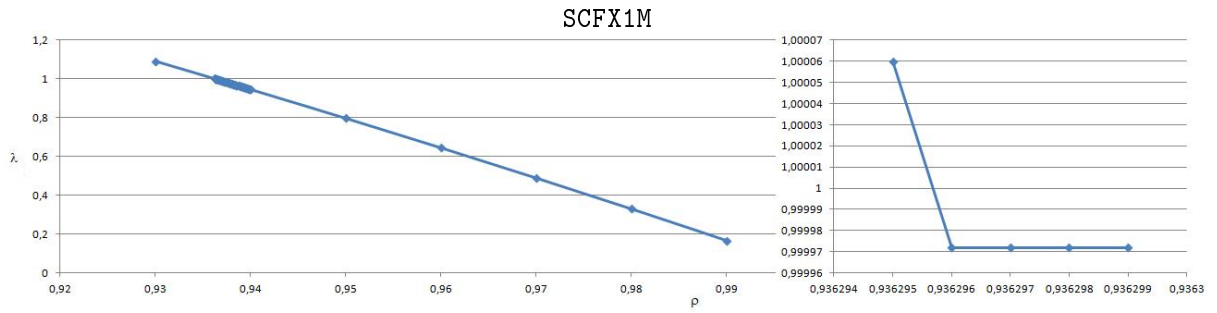
Apéndice



Apéndice



Apéndice



Bibliografía

- [1] F. Azadivar. Simulation optimization methodologies. In *Proceedings of the 31st conference on Winter simulation: Simulation—a bridge to the future-Volume 1*, pages 93–100. ACM, 1999.
- [2] A. Ben-Tal, S. Boyd, and A. Nemirovski. Extending scope of robust optimization: Comprehensive robust counterparts of uncertain problems. *Mathematical Programming*, 107(1):63–89, 2006.
- [3] A. Ben-Tal and A. Nemirovski. Robust convex optimization. *Mathematics of Operations Research*, 23(4):769–805, 1998.
- [4] A. Ben-Tal and A. Nemirovski. Robust solutions of linear programming problems contaminated with uncertain data. *Mathematical Programming*, 88(3):411–424, 2000.
- [5] A. Ben-Tal and A. Nemirovski. Robust optimization—methodology and applications. *Mathematical Programming*, 92(3):453–480, 2002.
- [6] D. Bertsimas and M. Sim. The price of robustness. *Operations Research*, 52(1):35–53, 2004.
- [7] J. Birge and F. Louveaux. *Introduction to stochastic programming*. Springer Verlag, 1997.
- [8] M. Canos Darós, M. Mocholí, V. Navarro, V. Sanchís, and R. Sala Garrido. Aplicaciones económicas de la optimización robusta. *Rect@*, 1997.
- [9] Y. Carson and A. Maria. Simulation optimization: methods and applications. In *Proceedings of the 29th conference on Winter simulation*, pages 118–126. IEEE Computer Society, 1997.
- [10] G. Dantzig. Maximization of a linear function of variables subject to linear inequalities. *New York*, 1951.
- [11] G. Dantzig. Linear programming under uncertainty. *Management Science*, 1(3/4):197–206, 1955.
- [12] J. Dupacová. Stochastic programming: Approximation via scenarios. 1996.
- [13] J. Dupacová, N. Gröwe-Kuska, and W. Römisch. Scenario reduction in stochastic programming. *Mathematical Programming*, 95(3):493–511, 2003.

- [14] L. El Ghaoui and H. Lebret. Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications*, 18:1035–1064, 1997.
- [15] L. El Ghaoui, F. Oustry, H. Lebret, et al. Robust solutions to uncertain semidefinite programs. *SIAM journal of optimization*, 9:33–52, 1998.
- [16] H. Frank. Optimum locations on a graph with probabilistic demands. *Operations Research*, 14(3):409–421, 1966.
- [17] F. Gould. Proximate Linear Programming: An Experimental Study of a Modified Simplex Algorithm for Solving Linear Programs with Inexact Data, 1971.
- [18] D. Heyman and M. Sobel. *Stochastic models in operations research, Vol. II: Stochastic optimization*, volume 2. Dover Pubns, 2003.
- [19] C. Jansson, D. Chaykin, and C. Keil. Rigorous error bounds for the optimal value in semidefinite programming. *Institute for Reliable Computing*, 500:1–4, 2007.
- [20] F. Ordonez and R. Freund. Computational experience and the explanatory value of condition measures for linear optimization. *SIAM Journal on Optimization*, 14(2):307–333, 2004.
- [21] J. Renegar. Some perturbation theory for linear programming. *Mathematical Programming*, 65(1):73–91, 1994.
- [22] A. Shapiro and A. Nemirovski. On complexity of stochastic programming problems. *Continuous optimization*, pages 111–146, 2005.
- [23] A. Soyster. Convex programming with set-inclusive constraints and applications to inexact linear programming. *Operations research*, 21(5):1154–1157, 1973.
- [24] J. Spall. Implementation of the simultaneous perturbation algorithm for stochastic optimization. *Aerospace and Electronic Systems, IEEE Transactions on*, 34(3):817–823, 1998.
- [25] J. Swisher, P. Hyden, S. Jacobson, and L. Schruben. A survey of simulation optimization techniques and procedures. In *Simulation Conference Proceedings, 2000. Winter*, volume 1, pages 119–128. IEEE, 2000.
- [26] G. Watson. Robust counterparts of errors-in-variables problems. *Computational Statistics & Data Analysis*, 52(2):1080–1089, 2007.