SciVerse ScienceDirect
Disponible en www.sciencedirect.com
Revista Iberoamericana de Automática e Informática industrial 10 (2013) 50–61
RIAI
www.elsevier.es/RIAI

Un algoritmo secuencial, aleatorio y óptimo para problemas de factibilidad robusta

T. Álamo[a], R. Tempo[b], D.R. Ramírez[a], A. Luque[a], E.F. Camacho[a]

[a]Dpto. Ingeniería de Sistemas y Automática, Escuela Técnica Superior de Ingenieros, Universidad de Sevilla, Camino Descubrimientos, s/n., 41092 Sevilla, España
[b]IEIIT-CNR, Politecnico di Torino, Corso Duca degli Abruzzi, 24, 10129 Torino, Italy

Resumen

En este trabajo (del cual se presentó una versión preliminar en Alamo et al. (2007)) se propone un algoritmo aleatorio para determinar la factibilidad robusta de un conjunto de desigualdades lineales matriciales (Linear Matrix Inequalities, LMI). El algoritmo está basado en la solución de una secuencia de problemas de optimización semidefinida sujetos a un bajo número de restricciones. Se aporta además una cota superior del número máximo de iteraciones que requiere el algoritmo para resolver el problema de factibilidad robusta. Finalmente, los resultados se ilustran mediante un ejemplo numérico. *Copyright © 2013 CEA. Publicado por Elsevier España, S.L. Todos los derechos reservados.*

Palabras Clave: factibilidad robusta, desigualdades lineales matriciales, algoritmos aleatorios, control robusto.

1. Introducción

Los problemas de factibilidad robusta constituyen una herramienta muy interesante para diferentes campos del control automático, principalmente en aquellos que manejan un paradigma de robustez ante incertidumbres de modelado. Gran parte de ese interés proviene del hecho de que ciertos problemas de diseño o análisis de controladores robustos pueden ser reescritos como problemas de factibilidad robusta. Este tipo de problemas implican la satisfacción de restricciones de análisis o diseño para todos los valores admisibles de los parámetros inciertos del modelo del sistema considerado. Usualmente el número de esas restricciones es infinito, debido a la posible variación de los parámetros inciertos, que pueden tomar cualquier valor de un conjunto de valores admisibles previamente especificado. Además de la propia satisfacción de las soluciones se persigue la optimización de una cierta función de coste, de manera que el controlador resultante de la resolución del problema de optimización será el mejor, según el criterio de funcionamiento elegido, de entre todos los que satisfagan las restricciones. Desde un punto de vista del formalismo matemático empleado, muchos de estos problemas de diseño y análisis robusto se resuelven como problemas de desigualdades matriciales lineales (LMI). Ejemplos de este tipo de problemas, los encontramos en el diseño de controladores robustos para sistemas con incertidumbres paramétricas Feron et al. (1996) o lineales de parámetros variables (LPV) (Becker and Packard (1994); Apkarian et al. (1995)), en la síntesis de filtros robustos (Tuan et al. (2001)) y en el análisis μ (Apkarian and Tuan (2000)) por citar algunos. También aparecen problemas de factibilidad robusta de LMIs en control predictivo basado en modelos (Wan and Kothare (2002)), siendo estos problemas muy interesantes debido a que la estabilidad robusta viene determinada por la factibilidad de las restricciones consideradas (Scokaert et al. (1999)).

Como se ha comentado en el párrafo anterior, habitualmente el número de restricciones presentes en los problemas de factibilidad robusta es o muy grande o incluso infinito, siendo en muchos casos la complejidad computacional de estos problemas del tipo NP-dura (Nemirovskii (1993)). En los últimos años, los algoritmos aleatorios (Tempo et al. (2005); Calafiore et al. (2007, 2011); Alamo et al. (2009)) han despertado el interés de los investigadores en control, fundamentalmente por la posibilidad de usarlos para evitar esa naturaleza NP-dura. Estos algoritmos permiten obtener en tiempo polinomial una solución aproximada que satisface la mayor parte de las restricciones asociadas a posibles valores de la incertidumbre de un problema de robustez. Además, el número de restricciones del problema original que son violadas por la solución aproximada se puede hacer tan pequeño como se desee.

Una clase destacada en las estrategias aleatorias son las basadas en métodos gradenciales (Polyak and Tempo (2001); Calafiore and Polyak (2001); Fujisaki et al. (2003); Liberzon and Tempo (2004)). Estas estrategias son capaces de encontrar, con probabilidad uno, una solución a un problema de control robusto basado en LMIs en un número finito de iteraciones, si el problema es factible. Estos métodos están basados en un esquema

Correos electrónicos: alamo@cartuja.us.es (T. Álamo), roberto.tempo@polito.it (R. Tempo), danirr@cartuja.us.es (D.R. Ramírez), amalia@cartuja.us.es (A. Luque), eduardo@esi.us.es (E.F. Camacho)

http://dx.doi.org/10.1016/j.riai.2012.11.005

iterativo, en el que la solución actual se actualiza en la dirección obtenida del gradiente aleatorio de una función de factibilidad apropiada.

Otro tipo importante de algoritmos aleatorios son los basados en versiones probabilísticas de los métodos de localización. La convergencia a la solución en estos métodos es, teóricamente, mejor que los gradenciales. Entre estos métodos cabe citar el algoritmo probabilístico del elipsoide (Oishi (2003); Kanev et al. (2003)) y la versión probabilística del método del plano de corte por el centro analítico de un conjunto de LMIs (Calafiore and Dabbene (2007)).

La estrategia del «escenario» juega también un papel importante en la resolución de problemas de optimización robusta. Tal y como se demuestra en Calafiore and Campi (2006), muestreando de manera apropiada el conjunto de restricciones, se obtiene un problema convexo (el «escenario») cuya solución es aproximadamente factible para el problema original (el cual implica un número mucho mayor, posiblemente infinito, de restricciones). Al incrementar el número de muestras, el número de restricciones del problema original que se violan tiende a cero.

La estrategia de «escenario», los métodos gradenciales y el método del elipsoide tienen una naturaleza muy diferente. La primera obtiene una solución aproximada mediante la resolución de un único problema convexo con un número alto de restricciones. Por otra parte los métodos gradenciales y el método del elipsoide obtienen una solución aproximada de manera secuencial, empleando un número considerable de iteraciones, en cada una de las cuales se actualiza una solución candidata mediante una sencilla regla sin optimización alguna.

En este trabajo se presenta un algoritmo aleatorio que resuelve el problema de obtener una solución robusta factible para un conjunto posiblemente infinito de LMIs. Este algoritmo no pertenece a ninguna de las clases anteriores y converge en un número finito de iteraciones, siendo además capaz de determinar la no factibilidad del problema. La solución se alcanza mediante una secuencia de problemas de optimización relativamente simples. El método propuesto comparte con los métodos gradenciales y de elipsoide su naturaleza secuencial. A diferencia de éstos utiliza una regla de actualización de la solución candidata basada en un problema de optimización en el que se emplea un reducido número de restricciones obtenidas del problema original. Una de las ventajas del método propuesto es que es capaz de determinar la no factibilidad del problema de factibilidad robusta. En la práctica el algoritmo funciona de manera muy satisfactoria, obteniendo una solución aproximadamente factible o detectando la no factibilidad en un número razonable de iteraciones.

El resto del trabajo está organizado como sigue: la sección 2 establece la clase de problemas que se considera. El concepto de corte normalizado válido se introduce en la sección 3. Por otra parte, la sección 4 presenta resultados preliminares, necesarios para demostrar la convergencia del algoritmo propuesto, el cual se presenta en la sección 5. La sección 6 muestra como resolver el problema de comprobar la factibilidad robusta de una solución candidata. La sección 7 presenta un ejemplo numérico, terminando el trabajo con una sección dedicada a conclusiones y los apéndices A y B dedicados a propiedades auxiliares y demostraciones.

2. Tipo de problema considerado

En este trabajo[1] se trata la solución del siguiente problema LMI robusto:

$$\begin{array}{ll} \textbf{encontrar} & \mathbf{z} \in \mathbb{R}^m \\ \text{s.a.} & A(\mathbf{z}, w) < 0, \quad \forall w \in W \end{array} \tag{1}$$

donde W es un conjunto compacto, $A(\mathbf{z}, w) = \sum_{i=1}^{m} z_i A_i(w)$, y $A_i(w) = A_i^\top(w) \in \mathbb{R}^{q\times q}$, $i = 1, \ldots, m$, $\forall w \in W$. Se asume que las componentes de las matrices $A_i(w)$, $i = 1, \ldots, m$ están acotadas para cada $w \in W$. Si $\mathbf{z}$ satisface que $A(\mathbf{z}, w) < 0$, $\forall w \in W$ entonces se dice que $\mathbf{z}$ es una solución factible robusta.

El conjunto de soluciones factibles de (1) se denota por $\mathcal{D}$,

$$\mathcal{D} = \{\ \mathbf{z} \in \mathbb{R}^m\ :\ A(\mathbf{z}, w) < 0, \quad \forall w \in W\ \}.$$

Nótese que la dependencia de $A(\mathbf{z}, w)$ con respecto a $\mathbf{z}$ es lineal. Esto implica que $A(\mu\mathbf{z}, w) = \mu A(\mathbf{z}, w)$, $\forall \mu \in \mathbb{R}$. De esto se deduce que si $\mathbf{z}$ es una solución factible robusta, entonces $\bar{\mathbf{z}} = \frac{\mathbf{z}}{\sqrt{\mathbf{z}^\top\mathbf{z}}}$ también lo es y $\bar{\mathbf{z}}^\top\bar{\mathbf{z}} = 1$. Por tanto, para estudiar la factibilidad robusta del problema (1), es suficiente con analizar si existe una solución factible robusta en la esfera unitaria $\{\ \mathbf{z} \in \mathbb{R}^m\ :\ \mathbf{z}^\top\mathbf{z} \le 1\ \}$. Esto motiva la noción de ϵ-factibilidad.

Definición 1. *Para un $\epsilon > 0$ dado, se dice que el problema de factibilidad robusta (1) es ϵ-factible si existe $\mathbf{z} \in \mathbb{R}^m$ tal que*

$$\begin{array}{rcl} A(\mathbf{z}, w) & \le & -\epsilon \mathrm{I}, \quad \forall w \in W \\ \mathbf{z}^\top\mathbf{z} & \le & 1. \end{array}$$

A continuación se verá como, aunque la dependencia con respecto a las variables de decisión en las desigualdades matriciales sea *afín*, este tipo de problema se puede reformular usando LMIs robustas en las que la dependencia es lineal. Considérese el problema de encontrar $\mathbf{x} \in \mathbb{R}^n$ tal que

$$F(\mathbf{x}, w) < 0, \quad \forall w \in W \tag{2}$$

con

$$F(\mathbf{x}, w) = F_0(w) + \sum_{i=1}^{n} x_i F_i(w)$$

y donde $F_i(w) = F_i^\top(w)$, $i = 0, \ldots, n$, $\forall w \in W$. Nótese que esta clase de problemas de factibilidad robusta aparece frecuentemente en control robusto (Boyd et al. (1994)).

Supóngase un valor $\mu > 0$. Entonces, teniendo en cuenta las restricciones dadas por la ecuación (2) se cumple que:

$$\mu F_0(w) + \sum_{i=1}^{n} \mu x_i F_i(w) < 0, \ \forall w \in W.$$

[1] La notación adoptada en este trabajo es como sigue: dado el vector $\mathbf{x}$, x_i denota la iésima componente y $\|\cdot\|_2$ la norma euclídea. Dada una matriz simétrica A, $\bar{\lambda}(A)$ denota el mayor de los autovalores; $A < 0$ denota que es definida negativa; dadas dos matrices simétricas A y B, $A < B$ denota que $A - B$ es definida negativa. Sea $x \in \mathbb{R}$, $x > 0$, $\lceil x \rceil$ denota el menor entero mayor o igual a x.

Denótese ahora $\mu x_i = z_i$, $i = 1, \ldots, n$ y $\mu = z_{n+1}$ y considérese el siguiente problema de factibilidad robusta de LMIs:

$$\textbf{encontrar } \mathbf{z} \in \mathbb{R}^{n+1} \textbf{ tal que } \forall w \in W:$$
$$A(\mathbf{z}, w) = \begin{bmatrix} z_{n+1}F_0(w) + \sum_{i=1}^{n} z_i F_i(w) & 0 \\ 0 & -z_{n+1} \end{bmatrix} < 0. \quad (3)$$

Los problemas de factibilidad (2) y (3) son equivalentes: si $\mathbf{x} \in \mathbb{R}^n$ es una solución factible del problema (2) entonces $\mathbf{z} = [\mathbf{x}^\top \ 1]^\top$ es, a su vez, solución factible del problema (3). De manera análoga, si $\mathbf{z}$ es una solución factible de (3) entonces $z_{n+1} > 0$ y $\mathbf{x} = \frac{1}{z_{n+1}}[z_1 \ z_2 \ \ldots \ z_n]^\top$ es una solución factible del problema (2). Por tanto, el problema de factibilidad (3) pertenece a la clase de problemas de factibilidad dados por (1). Es decir, aunque la dependencia con respecto a la variable de decisión sea afín, el problema se puede reescribir de manera que la dependencia sea lineal.

Finalmente, tal y como se enuncia en el siguiente teorema, el problema de factibilidad (1) se puede formular como un problema de optimización.

Teorema 1. *Denótese por γ^* la solución del siguiente problema de minimización:*

$$\begin{aligned} \gamma^* = \ & \min_{\mathbf{z},\gamma} \ \gamma \\ & s.a. \ A(\mathbf{z}, w) \leq \gamma \mathrm{I}, \ \ \forall w \in W \\ & \ \mathbf{z}^\top \mathbf{z} \leq 1. \end{aligned} \quad (4)$$

Entonces:

(I) *El problema (1) es factible si y sólo si $\gamma^* \neq 0$.*
(II) *El problema (1) es no ϵ-factible para cada $\epsilon > -\gamma^*$.*

Prueba: Véase el apéndice B.

Por tanto, se puede resolver el problema de factibilidad considerado como si fuese un problema de optimización.

3. Cortes normalizados válidos

En esta sección se presenta el concepto de cortes normalizados válidos. Este concepto es fundamental porque, como se mostrará en esta sección, estos cortes normalizados válidos permiten obtener cotas inferiores de γ^*.

Definición 2. *Sea $\mathbf{c} \in \mathbb{R}^m$. Se dice que $\mathbf{c}$ es un corte normalizado válido de $A(\mathbf{z}, w)$ si existe $v \in \mathbb{R}^q$ tal que $v^\top v = 1$ y $c_i = v^\top A_i(w) v$, $i = 1, \ldots, m$.*

Dado un $w \in W$,

$$\bar{\lambda}(A(\mathbf{z}, w)) = \max_{v \in \mathbb{R}^q, v^\top v = 1} v^\top A(\mathbf{z}, w) v = \max_{v \in \mathbb{R}^q, v^\top v = 1} \sum_{i=1}^{m} (v^\top A_i(w) v) z_i.$$

De ahí se infiere la siguiente propiedad.

Propiedad 1. *Denótese*

$$C(w) = \{\mathbf{c} \in \mathbb{R}^m \ : \ \textit{existe} \ v \in \mathbb{R}^q \ \textit{tal que}$$
$$v^\top v = 1 \ \textit{y} \ c_i = v^\top A_i(w) v, \ i = 1, \ldots, m\}.$$

Con esta notación, resulta que $A(\mathbf{z}, w) \leq \gamma \mathrm{I}$ si y sólo si se verifica que $\max_{\mathbf{c} \in C(w)} \mathbf{c}^\top \mathbf{z} \leq \gamma$.

La unión de todos los conjuntos $C(w)$ se denotará por C:

$$C = \bigcup_{w \in W} C(w).$$

A C se le denomina el conjunto de todos los cortes normalizados válidos. De esta definición y de la propiedad 1 resulta que (4) se puede reescribir como:

$$\begin{aligned} \gamma^* = \ & \min_{\mathbf{z},\gamma} \ \gamma \\ & s.a. \ \mathbf{c}^\top \mathbf{z} \leq \gamma, \ \ \forall \mathbf{c} \in C \\ & \ \mathbf{z}^\top \mathbf{z} \leq 1. \end{aligned} \quad (5)$$

La factibilidad de cualquier vector $\mathbf{z}$ se puede determinar por medio del conjunto de todos los cortes normalizados válidos:

Propiedad 2. *El vector $\mathbf{z}$ es una solución factible del problema (1) si y sólo si $\mathbf{c}^\top \mathbf{z} < 0$, $\forall \mathbf{c} \in C$.*

Prueba: La prueba se deriva directamente del hecho de que $A(\mathbf{z}, w)$ es definida negativa para todo $w \in W$ si y sólo si $\forall w \in W$ y $\forall v \ : \ v^\top v = 1$ se cumple que $v^\top \left(\sum_{i=1}^{m} A_i(w) z_i \right) v < 0$. De la definición de C, se infiere que la última desigualdad equivale a $\mathbf{c}^\top \mathbf{z} < 0$, $\forall \mathbf{c} \in C$. □

Definición 3. *El vector $\mathbf{h}$ pertenece a la envoltura convexa del conjunto C, (denotada por $\mathrm{Co}\{C\}$) si y sólo si existen $\mathbf{c}_1, \ldots, \mathbf{c}_p$ y $\lambda_1, \ldots, \lambda_p$ tales que*

$$\mathbf{h} = \sum_{i=1}^{p} \lambda_i \mathbf{c}_i \ \text{ y } \ 1 = \sum_{i=1}^{p} \lambda_i$$

siendo $\lambda_i \geq 0$, $i = 1, \ldots, p$ y $\mathbf{c}_i \in C$, $i = 1, \ldots, p$.

La siguiente propiedad establece que se puede inferir una cota inferior de γ^* directamente de todo elemento de $\mathrm{Co}\{C\}$.

Propiedad 3. *Se tiene que:*

$$\gamma^* \geq -\|\mathbf{h}\|_2, \quad \forall \mathbf{h} \in \mathrm{Co}\{C\}.$$

Prueba: Véase el apéndice B.

Es decir, a partir de los cortes normalizados válidos se pueden inferir cotas inferiores de la solución del problema (4), de ahí que el concepto de corte normalizado válido sea fundamental a la hora de formular un algoritmo para resolver el problema (4) . Esas cotas obtenidas de los cortes normalizados válidos pueden ser demasiado conservadoras, de manera que a continuación se verá cómo mejorarlas. Esto es necesario, porque el algoritmo propuesto deberá obtener una cota inferior de γ^* en cada iteración.

4. Obtención de una cota inferior de γ^*

En esta sección se demuestra que se puede obtener una cota inferior de la solución del problema de optimización (4), γ^*, por medio de la solución de otro problema de optimización.

Tal y como se indica en la propiedad 3, se puede obtener una cota inferior de γ^* directamente de cada elemento de $\mathrm{Co}\{C\}$.

Es decir, $\gamma^* \geq -\|\mathbf{h}\|_2$, $\forall \mathbf{h} \in \mathrm{Co}\{C\}$. Por otra parte, dado un $\mathbf{h} \in \mathrm{Co}\{C\}/0$, entonces $\hat{\mathbf{z}} = -\frac{\mathbf{h}}{\|\mathbf{h}\|_2}$ satisface $\hat{\mathbf{z}}^\top \hat{\mathbf{z}} \leq 1$. Por tanto, el par

$$(\hat{\mathbf{z}}, \hat{\gamma}) = \left(-\frac{\mathbf{h}}{\|\mathbf{h}\|_2}, -\|\mathbf{h}\|_2\right)$$

podría ser considerado una solución candidata del problema (4). Evidentemente esa solución candidata no sería una solución de (4) si existiese un $\hat{w} \in W$ tal que $\bar{\lambda}(A(\hat{\mathbf{z}}, \hat{w})) > 0$. En ese caso, se podría usar $\hat{w}$ para obtener una cota inferior mejorada de γ^*. Esto es lo que se establece en la siguiente propiedad.

Propiedad 4. *Considérese el siguiente problema*

$$\begin{array}{rll} \gamma_c^* = & \min\limits_{\mathbf{z},\gamma} & \gamma \\ & s.a. & \mathbf{h}^\top \mathbf{z} \leq \gamma, \\ & & A(\mathbf{z}, w) \leq \gamma \mathrm{I}, \quad \forall w \in S \\ & & \mathbf{z}^\top \mathbf{z} \leq 1. \end{array} \tag{6}$$

Supóngase que $S \subseteq W$, $\mathbf{h} \in \mathrm{Co}\{C\}$, $\mathbf{h} \neq 0$ *y que existe* $\hat{w} \in S$ *tal que* $\bar{\lambda}\left(A(\frac{-\mathbf{h}}{\sqrt{\mathbf{h}^\top \mathbf{h}}}, \hat{w})\right) > 0$. *Supóngase también que* $\max\limits_{\mathbf{c} \in C} \|\mathbf{c}\|_2 < \sigma$. *Entonces, denotando* $\hat{\gamma} = -\sqrt{\mathbf{h}^\top \mathbf{h}}$ *se cumple que:*

$$\gamma^* \geq \gamma_c^* \geq \hat{\gamma}\left(1 - \frac{\hat{\gamma}^2}{8\sigma^2}\right) \geq \hat{\gamma}.$$

Prueba: Véase el apéndice B.

De esta manera, resolviendo el problema (6) se obtiene una cota inferior de la solución que luego será utilizada como base para resolver el problema (4) mediante el algoritmo que se propone en la sección siguiente.

5. Algoritmo propuesto

En esta sección se presentará un algoritmo que resuelve el problema de factibilidad robusta (1). Considérese un entero dado, $N_{max} \geq 1$, y un número real $\epsilon \in (0, 1)$ (el papel de estos números en el funcionamiento del algoritmo se clarificará más adelante). El algoritmo propuesto se detalla a continuación.

Algoritmo 1. *Problema de factibilidad*

1. *Escójase un elemento de W, denotado por* w_0 *y resuélvase el problema de minimización:*

$$\begin{array}{rllll} \gamma_0^* = & \min\limits_{\mathbf{z},\gamma} & & \gamma & \\ & s.a. & A(\mathbf{z}, w_0) & \leq & \gamma \mathrm{I} \\ & & \mathbf{z}^\top \mathbf{z} & \leq & 1. \end{array} \tag{7}$$

   Denótese por $(\mathbf{z}_0^*, \gamma_0^*)$ *la solución de este problema de optimización. Sea* $\mathbf{h}_0 = \gamma_0^* \mathbf{z}_0^*$, $S_0 = \{w_0\}$ *y* $k = 0$.
2. *Compruébese si* $\mathbf{z}_k^*$ *es una solución factible robusta del problema (1) (esto puede hacerse de manera probabilística como se demuestra en la siguiente sección).*
3. *Si* $\mathbf{z}_k^*$ *es una solución robusta factible (en sentido probabilístico) entonces PARAR. Si no, obtener*[2] $w_{k+1} \in W$ *tal que* $\bar{\lambda}(A(\mathbf{z}_k^*, w_{k+1})) > 0$.
4. *Sea* $S_{k+1} = S_k \bigcup w_{k+1}$. *Si el número de elementos de* S_{k+1} *es mayor que* N_{max}, *entonces eliminar de* S_{k+1} *el vector* $\hat{w} \in S_{k+1}$ *que minimiza* $\bar{\lambda}(A(\mathbf{z}_k^*, w))$.
5. *Resolver el problema de optimización*

$$\begin{array}{rll} \gamma_{k+1}^* = & \min\limits_{\mathbf{z},\gamma} & \gamma \\ & s.a. & \mathbf{h}_k^\top \mathbf{z} \leq \gamma, \\ & & A(\mathbf{z}, w) \leq \gamma \mathrm{I}, \quad \forall w \in S_{k+1} \\ & & \mathbf{z}^\top \mathbf{z} \leq 1. \end{array} \tag{8}$$

   Denótese por $(\mathbf{z}_{k+1}^*, \gamma_{k+1}^*)$ *su solución. Sea* $\mathbf{h}_{k+1} = \gamma_{k+1}^* \mathbf{z}_{k+1}^*$.
6. *Si* $\gamma_{k+1}^* > -\epsilon$ *entonces el problema de factibilidad robusta (1) no es* ϵ*-factible. PARAR. Si no, hacer* $k = k + 1$ *e ir al paso 2.*

Comentario 1. *El problema que se resuelve en el paso 1 para obtener una solución inicial es diferente (falta la restricción* $\mathbf{h}_k^\top \mathbf{z} \leq \gamma$*) al que luego se resuelve en cada iteración en el paso 5. La razón es que inicialmente no se cuenta con el vector* $\mathbf{h}_k$ *correspondiente. Sin embargo, el valor obtenido* γ_0^* *es también una cota inferior de* γ^* *(véase la propiedad 6 en el apéndice A).*

El siguiente teorema afirma que el algoritmo encuentra una solución factible (probabilísticamente) o determina la no ϵ-factibilidad del problema en un número finito de iteraciones. La factibilidad será determinista o probabilística dependiendo de como se haya implementado el paso 2 del algoritmo. Esto se desarrollará en la sección 6.

Teorema 2. *Supóngase que* $\max\limits_{\mathbf{c} \in C} \|\mathbf{c}\|_2 < \sigma$ *y* $N_{max} \geq 1$. *Entonces, el algoritmo propuesto encuentra una solución factible robusta (probabilística) del problema (1), o determina que es no* ϵ*-factible en como máximo* k_{max} *iteraciones, siendo*

$$k_{max} = \left\lceil \frac{4\sigma^2}{\epsilon^2} \right\rceil + \left\lceil \left(\frac{4\sigma^2}{\epsilon^2}\right) \ln\left(\ln\left(\frac{\sigma^2}{\epsilon^2}\right)\right) \right\rceil.$$

Prueba: En primer lugar, se probará que $\mathbf{h}_k \in \mathrm{Co}\{C\}$, $k = 0, 1, \ldots$. De hecho, la propiedad 6 (véase el apéndice A) garantiza que $\mathbf{h}_0 \in \mathrm{Co}\{C\}$. Supóngase ahora que $\mathbf{h}_k \in \mathrm{Co}\{C\}$. Por construcción, $\mathbf{h}_{k+1} = \gamma_{k+1}^* \mathbf{z}_{k+1}^*$ donde $(\mathbf{z}_{k+1}^*, \gamma_{k+1}^*)$ es la solución óptima del problema

$$\begin{array}{rll} \gamma_{k+1}^* = & \min\limits_{\mathbf{z},\gamma} & \gamma \\ & s.a. & \mathbf{h}_k^\top \mathbf{z} \leq \gamma, \\ & & A(\mathbf{z}, w) \leq \gamma \mathrm{I}, \quad \forall w \in S_{k+1} \\ & & \mathbf{z}^\top \mathbf{z} \leq 1. \end{array} \tag{9}$$

Como $\mathbf{h}_k \in \mathrm{Co}\{C\}$ y $S_{k+1} \subseteq W$, y por la propiedad 5 (véase el apéndice A) se infiere que $\mathbf{h}_{k+1} \in \mathrm{Co}\{C\}$. Por tanto, $\mathbf{h}_0 \in \mathrm{Co}\{C\}$ y $\mathbf{h}_k \in \mathrm{Co}\{C\}$ lo que implica que $\mathbf{h}_{k+1} \in \mathrm{Co}\{C\}$. Esto significa que

$$\mathbf{h}_k \in \mathrm{Co}\{C\}, \quad k = 0, 1, \ldots \tag{10}$$

Teniendo en cuenta (10) y que $S_k \subseteq W$, $k = 0, 1, \ldots$, se infiere de las propiedades 5 y 6 que

$$\gamma^* \geq \gamma_k^*, \quad k = 0, 1, \ldots$$

Por tanto, si $\gamma_k^* > -\epsilon$ entonces $\gamma^* \geq \gamma_k^* > -\epsilon$, lo que (de acuerdo con el teorema 1) implica que el problema no es ϵ-factible. Esto significa que el algoritmo es capaz de clasificar

[2] La forma de obtener $w(k+1)$ no es algo crítico, y puede hacerse por ejemplo escogiendo aquel que resulta no factible por un mayor margen, o también el primero que se encuentra al evaluar las restricciones.

correctamente los problemas no ϵ-factibles como tales, lo que prueba que el algoritmo es correcto.

Supóngase que tras $k_{\text{máx}}$ iteraciones, el algoritmo no ha encontrado una solución factible robusta (probabilística) ni ha determinado que el problema es no ϵ-factible. Entonces,

$$\gamma_k^* \leq -\epsilon < 0, \quad k = 0, \ldots, k_{max}. \tag{11}$$

De $\gamma_k^* < 0$, $k = 0, \ldots, k_{max}$, y las propiedades 5 y 6 se concluye que $\|\mathbf{z}_k^*\|_2 = 1$, $k = 0, \ldots, k_{max}$. Por tanto, de las igualdades $\mathbf{h}_k = \gamma_k^* \mathbf{z}_k^*$ and $\|\mathbf{z}_k^*\|_2 = 1$ se infiere que $\gamma_k^* = -\sqrt{\mathbf{h}_k^\top \mathbf{h}_k}$. Es decir,

$$\mathbf{z}_k^* = \frac{\mathbf{h}_k}{\gamma_k} = \frac{-\mathbf{h}_k}{\sqrt{\mathbf{h}_k^\top \mathbf{h}_k}}, \quad k = 0, \ldots, k_{max}.$$

Por construcción, $w_{k+1} \in S_{k+1}$ y

$$\bar{\lambda}\left(A\left(\mathbf{z}_k^*, w_{k+1}\right)\right) = \bar{\lambda}\left(A\left(\frac{-\mathbf{h}_k}{\sqrt{\mathbf{h}_k^\top \mathbf{h}_k}}, w_{k+1}\right)\right) > 0.$$

Teniendo en cuenta la propiedad 4 y la definición de γ_{k+1}^* en el paso 5 del algoritmo, se concluye que

$$\gamma_{k+1}^* \geq \gamma_k^* \left(1 - \frac{(\gamma_k^*)^2}{8\sigma^2}\right) \geq \gamma_k^*. \tag{12}$$

Denótese ahora $\Gamma_k = -\gamma_k^*$ y $\tau = \frac{1}{8\sigma^2}$. Con esta notación se puede reescribir la ecuación (12) como

$$0 \leq \Gamma_{k+1} \leq \Gamma_k \left(1 - \tau\Gamma_k^2\right). \tag{13}$$

Nótese que $\Gamma_0 = -\gamma_k^* = \|\mathbf{h}_0\|_2$. Debido a la convexidad de la norma $\|\cdot\|_2$, se concluye que $\max\limits_{\mathbf{h}\in \mathrm{Co}\{C\}} \|\mathbf{h}\|_2 = \max\limits_{\mathbf{c}\in C} \|\mathbf{c}\|_2 < \sigma$. Recordando que $\mathbf{h}_0 \in \mathrm{Co}\{C\}$ se infiere que

$$\Gamma_0 < \sigma. \tag{14}$$

Por otra parte la suposición de que $\gamma_k^* \leq -\epsilon$, $k = 0, \ldots, k_{max}$ implica que $\epsilon \leq -\gamma_0^* = \Gamma_0 < \sigma$. Por tanto,

$$\sigma > \epsilon > 0. \tag{15}$$

De $\sigma > \epsilon$ se concluye también que

$$0 < \tau\epsilon^2 = \frac{\epsilon^2}{8\sigma^2} < 1. \tag{16}$$

De las ecuaciones (13), (14), (15) y de (16) y la propiedad 10 (véase el apéndice A) se infiere que

$$\Gamma_k < \epsilon \text{ if } k \geq \left\lceil \frac{1}{2\tau\epsilon^2} \right\rceil + \left\lceil \frac{\ln\ln \frac{\sigma^2}{\epsilon^2}}{2\tau\epsilon^2} \right\rceil.$$

Recuérdese que $\Gamma_k^* = -\gamma_k^*$ y que $\tau = \frac{1}{8\sigma^2}$. Entonces, $\gamma_k^* > -\epsilon$ si

$$k \geq \left\lceil \frac{4\sigma^2}{\epsilon^2} \right\rceil + \left\lceil \left(\frac{4\sigma^2}{\epsilon^2}\right) \ln\ln \frac{\sigma^2}{\epsilon^2} \right\rceil = k_{max}.$$

Esto contradice (11), lo que prueba que el algoritmo para en no más de k_{max} iteraciones. □

Comentario 2. *Nótese que el algoritmo encuentra una solución o determina la no ϵ-factibilidad para todo $N_{max} \geq 1$ en como máximo k_{max} iteraciones. Sin embargo, la experiencia demuestra que el número de iteraciones requeridas decrece notablemente al incrementar N_{max} (véase la sección 7).*

5.1. *Acotación del valor de σ*

Recuérdese que cada conjunto $C(w)$ definido en la propiedad 1, es compacto. Esto y la compacidad de W implica que $C = \sum\limits_{w\in W} C(w)$ es un conjunto compacto. De ahí y de la suposición de que las componentes de cada matriz $A_i(w)$, $i = 1, \ldots, m$ están acotadas para todo $w \in W$ se concluye que existe una constante σ tal que $\|\mathbf{c}\|_2 \leq \sigma$, para cada $\mathbf{c} \in C$. Por tanto, la suposición de la existencia de σ no es restrictiva.

El algoritmo propuesto se puede implementar sin obtener una cota de $\max\limits_{\mathbf{c}\in C} \|\mathbf{c}\|_2$. Sin embargo, como el número máximo de iteraciones depende de esa cota, puede ser interesante calcular σ tal que $\max\limits_{\mathbf{c}\in C} \|\mathbf{c}\|_2 < \sigma$. Así se podría saber por anticipado cuantas iteraciones se requerirían, como máximo, para resolver el problema.

Nótese que

$$\begin{aligned} |c_i| &\leq \max_{w\in W} \max_{v^\top v=1} |v^\top A_i(w) v| \\ &= \max_{w\in W} \sqrt{\bar{\lambda}(A_i^2(w))} \\ &= \max_{w\in W} \bar{\sigma}(A_i(w)), \quad i = 1\ldots, m \ \ \forall \mathbf{c} \in C \end{aligned}$$

donde $\bar{\sigma}(A_i(w))$ denota el mayor de los valores singulares de la matriz simétrica $A_i(w)$.

Cuando la dependencia de las matrices $A_i(w)$ respecto de w es afín (o lineal-fraccional) y W es un conjunto estructurado y acotado por una norma ($W = \{\, \Delta \in \mathbf{\Delta} \;:\; \|\Delta\| \leq \rho \,\}$), entonces se pueden usar los resultados de la teoría μ (Zhou et al. (1996)) o técnicas basadas en relajaciones para obtener $\sigma_1, \sigma_2, \ldots, \sigma_m$ tales que

$$\max_{w\in W} \bar{\sigma}(A_i(w)) < \sigma_i, \quad i = 1, \ldots, m.$$

Así, se obtendría la siguiente cota conservadora

$$\max_{\mathbf{c}\in C} \|\mathbf{c}\|_2 < \sqrt{\sum_{i=1}^{m} \sigma_i^2}$$

6. Comprobación de la factibilidad robusta

Es posible decidir, de manera determinista, si una solución candidata es factible robusta usando resultados de la teoría μ (Zhou et al. (1996)), y técnicas basadas en relajaciones (Ben-Tal and Nemirovski (2001)). Sin embargo, estos resultados y técnicas son muy conservadores. En algunos casos significativos (por ejemplo, cuando existe dependencia afín con respecto a w), se pueden utilizar resultados basados en vértices (Barmish (1994); Boyd et al. (1994)). En estos casos, es suficiente con comprobar la factibilidad para un número finito de elementos de W, que forman el conjunto W_F. Si el número de elementos de W_F es razonable, el paso 2 del algoritmo propuesto se puede implementar de manera determinista. En ese caso existen diversas formas de visitar cada uno de los elementos de W_F, por ejemplo la «función de planificación» presentada en Liberzon and Tempo (2004).

Sin embargo, en la mayoría de los resultados basados en vértices, el número de elementos en W_F crece de manera exponencial con la dimensión de W (debido a que la mayor parte de los problemas de factibilidad robusta son NP-duros). Para agravar la situación, en muchos problemas de este tipo no es posible utilizar los resultados basados en vértices, debiéndose utilizar relajaciones del concepto de factibilidad (Barmish and Scherbakov (2000)).

Definición 4. *Considérese una medida de probabilidad «Prob» sobre un conjunto incierto W. El vector $\mathbf{z} \in \mathbb{R}^m$ se dice que es una solución factible robusta de nivel δ si*

$$\text{Prob}\ \{\ w \in W\ :\ \bar{\lambda}(A(\mathbf{z}, w)) \geq 0\ \} \leq \delta.$$

El siguiente teorema afirma que se puede usar el algoritmo propuesto para obtener una solución factible robusta de nivel δ con una probabilidad de fallo β pre-especificada.

Teorema 3. *Supóngase que dado $0 < \beta < 1$ y $0 < \delta < 1$, se sustituyen los pasos 2 y 3 del algoritmo propuesto por*

2. *Obtener el menor entero N_k tal que $(1-\delta)^{N_k} < \frac{6\beta}{\pi^2(1+k)^2}$. Escoger, de acuerdo con la medida de probabilidad «Prob», N_k elementos de W: $\bar{w}_1, \ldots, \bar{w}_{N_k}$.*
3. *Si $\bar{\lambda}(A(\mathbf{z}_k^*, \bar{w}_i)) < 0$, $i = 1 \ldots, N_k$, entonces clasificar $\mathbf{z}_k^*$ como una solución factible robusta de nivel δ. PARAR. Si no, hacer w_{k+1} igual a uno de los elementos de $\{\ \bar{w}_i\ :\ \bar{\lambda}(A(\mathbf{z}_k^*, \bar{w}_i)) \geq 0, i = 1 \ldots, N_k\ \}$.*

Con estas modificaciones, si el algoritmo propuesto termina clasificando a un $\mathbf{z}_k^*$ dado como una solución factible robusta de nivel δ, entonces $\mathbf{z}_k^*$ es una solución factible robusta de nivel δ con probabilidad mayor o igual a $1-\beta$.

PRUEBA: Véase el apéndice B.

7. Un ejemplo numérico

Considérese el siguiente sistema

$$\begin{cases} \dot{x} &= Ax + Bu + Ew \\ y &= Cx + Du \end{cases}$$

donde $A \in \mathbb{R}^{n_x \times n_x}$, $B \in \mathbb{R}^{n_x \times n_u}$, $E \in \mathbb{R}^{n_x \times n_w}$, $C \in \mathbb{R}^{n_y \times n_x}$ y $D \in \mathbb{R}^{n_y \times n_u}$ están afectadas por incertidumbre intervalar. Esto es,

$$\begin{aligned} A \in \mathcal{A}_I &= \{\ A\ :\ |A_{i,j} - \tilde{A}_{i,j}| \leq 0{,}02, \forall i, \forall j\ \} \\ B \in \mathcal{B}_I &= \{\ B\ :\ |B_{i,j} - \tilde{B}_{i,j}| \leq 0{,}02, \forall i, \forall j\ \} \\ E \in \mathcal{E}_I &= \{\ E\ :\ |E_{i,j} - \tilde{E}_{i,j}| \leq 0{,}02, \forall i, \forall j\ \} \\ C \in C_I &= \{\ C\ :\ |C_{i,j} - \tilde{C}_{i,j}| \leq 0{,}02, \forall i, \forall j\ \} \\ D \in \mathcal{D}_I &= \{\ D\ :\ |D_{i,j} - \tilde{D}_{i,j}| \leq 0{,}02, \forall i, \forall j\ \} \end{aligned}$$

donde

$$\tilde{A} = \begin{bmatrix} 0{,}083 & 0{,}650 & -0{,}627 & -0{,}624 & 0{,}091 \\ -0{,}832 & -0{,}261 & -0{,}023 & 0{,}275 & 0{,}658 \\ -0{,}985 & -0{,}447 & -0{,}007 & 0{,}731 & -0{,}942 \\ 0{,}179 & -0{,}957 & 0{,}649 & -0{,}873 & -0{,}350 \\ -0{,}305 & 0{,}239 & 0{,}610 & -0{,}752 & -0{,}624 \end{bmatrix} \tag{17}$$

$$\tilde{B} = \begin{bmatrix} 0{,}316 & 0{,}545 \\ 0{,}819 & -0{,}355 \\ 0{,}303 & 0{,}120 \\ -0{,}518 & -0{,}490 \\ -0{,}889 & -0{,}955 \end{bmatrix}, \quad \tilde{E} = \begin{bmatrix} -0{,}935 & 0{,}282 \\ 0{,}023 & -0{,}753 \\ -0{,}597 & 0{,}329 \\ 0{,}152 & -0{,}010 \\ -0{,}680 & 0{,}858 \end{bmatrix} \tag{18}$$

$$\tilde{C} = \begin{bmatrix} -0{,}433 & 0{,}131 & 0{,}877 & 0{,}986 & 0{,}179 \\ 0{,}823 & 0{,}551 & -0{,}664 & -0{,}930 & -0{,}747 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} \tag{19}$$

$$\tilde{D} = \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 0{,}876 & -0{,}368 \\ 0{,}317 & 0{,}773 \end{bmatrix}. \tag{20}$$

El número total de parámetros inciertos es $n_x(n_x + n_u + n_w) + n_y(n_x + n_u) = 73$. Esto significa que la incertidumbre se puede parametrizar mediante un vector de 73 componentes: $W = \{\ w \in \mathbb{R}^{73}\ :\ \|w\|_\infty \leq 0{,}02\ \}$.

El objetivo del problema es encontrar una matriz de ganancias de realimentación del vector de estados K de manera que se minimice la ganancia L_2 del sistema. Se sabe (Boyd et al. (1994)) que la ganancia L_2 de un sistema con incertidumbres en bucle cerrado está acotada por φ si existe una matriz $P > 0$ tal que

$$\frac{d}{dt}(x^\top Px) + y^\top y - \varphi^2 w^\top w \leq 0.$$

Denotando $Q = P^{-1}$ y $Y = KQ$, este problema de síntesis se puede reescribir como

$$\min_{Q,Y,\varphi} \quad \varphi^2$$
$$s.a.$$
$$\begin{bmatrix} AQ + QA^\top + BY + Y^\top B^\top & * & * & * \\ E^\top & -\varphi^2 \mathrm{I} & * & * \\ CQ + DY & 0 & -\mathrm{I} & * \\ 0 & 0 & 0 & -Q \end{bmatrix} < 0 \tag{21}$$
$$\forall A \in \mathcal{A}_I, \forall B \in \mathcal{B}_I, \forall C \in C_I, \forall D \in \mathcal{D}_I$$

Debido a la naturaleza intervalar de la incertidumbre y la dependencia afín de la misma, es suficiente con comprobar la factibilidad de las realizaciones extremas de la incertidumbre. Esto es, hay que comprobar los $2^{73} > 10^{21}$ vertices del hipercubo W (Boyd et al. (1994)). A fin de reducir el número de vértices, se pueden utilizar los resultados de Alamo et al. (2008), que en este caso reducen el número de vértices a $2^{18} = 262{,}144$.

Usando un método de bisección en la variable de decisión φ, este problema de optimización puede resolverse usando una secuencia de problemas de factibilidad robusta de la forma: Dado $\varphi > 0$, obtener (si es posible) Q y Y tal que la restricción robusta (21) se satisfaga para cada posible $w \in W$. Teniendo en cuenta la simetría de Q, es posible parametrizar las matrices $Q \in \mathbb{R}^{n_x \times n_x}$ y $Y \in \mathbb{R}^{n_u \times n_x}$ mediante un vector de 25 componentes. Por tanto, el problema de factibilidad anterior se puede reescribir como un problema de factibilidad de la forma:

$$\begin{aligned} &\textbf{encontrar} \quad \mathbf{x} \in \mathbb{R}^{25} \\ &\text{s.a.} \qquad F_0(w) + \sum_{i=1}^{25} x_i F_i(w) < 0, \quad \forall w \in W. \end{aligned}$$

Usando la estrategia de la sección 2, se puede reescribir este problema como un problema de factibilidad de la forma:

$$\begin{aligned} &\textbf{encontrar} \quad \mathbf{z} \in \mathbb{R}^{26} \\ &\text{s.a.} \qquad A(\mathbf{z}, w) < 0, \quad \forall w \in W \end{aligned}$$

donde $A(\mathbf{z}, w) = \sum_{i=1}^{26} z_i A_i(w)$.

Dadas las dimensiones de este ejemplo podemos comprobar si un par (Q, Y) dado es una solución robusta factible (para un φ dado). Esto significa que dado $\mathbf{z}$, se puede comprobar de manera determinista si $A(\mathbf{z}, w) < 0$, $\forall w \in W$. Por tanto, cada uno de los problemas de factibilidad requeridos para obtener un valor óptimo de φ puede resolverse usando el algoritmo propuesto sin utilizar la noción de factibilidad de nivel δ (es decir, el paso 2 del algoritmo se ha implementado de manera determinista).

Nótese que cuando el algoritmo de bisección está cerca del valor óptimo de φ, el problema de factibilidad asociado será ϵ-factible (si es que es factible) sólo para valores pequeños de ϵ. Esto tiene dos consecuencias:

1. El valor óptimo obtenido de φ dependerá de la elección de ϵ. Cuanto más pequeño sea ϵ, más cerca estará el valor de φ de la solución real del problema de optimización robusta de la ganancia L_2.
2. En el método de bisección usado para calcular el menor valor de φ para el que el problema de factibilidad robusta asociado es ϵ-factible, el algoritmo propuesto se puede encontrar las dos siguientes situaciones cuando se aproxime al valor final de φ:
   - El problema es ϵ-factible, pero al acercarse φ a la solución final, el radio de factibilidad del problema puede ser muy pequeño. Este radio de factibilidad reducido incrementará el número de iteraciones necesarias para obtener una solución factible.
   - El problema es no ϵ-factible, pero existe $\tilde{\epsilon}$, sólo ligeramente menor que ϵ, de manera que el problema es $\tilde{\epsilon}$-factible. El algoritmo tiene que determinar la no ϵ-factibilidad de un problema que es «casi» ϵ-factible. Por tanto también se incrementa el número de iteraciones requeridas para determinar la no ϵ-factibilidad del problema.

Lo anterior supone que el ejemplo que se considera aquí es muy apropiado para comprobar las propiedades de un algoritmo diseñado para el problema de factibilidad robusta.

A fin de ilustrar los efectos de ϵ y N_{max} en el rendimiento del algoritmo propuesto, se ha resuelto el problema de optimización robusta de la ganancia L_2 para diversos valores del par (N_{max}, ϵ). El algoritmo ha sido implementado en Matlab y ejecutado en un ordenador portátil con sistema operativo Windows XP y procesador Intel Pentium M750 a 1.86 Ghz. La tabla 1 muestra los resultados para $\epsilon = 0{,}01$ y diferentes valores de N_{max}. En dicha tabla, *max iteraciones* significa que todos los problemas de factibilidad de la secuencia (correspondiente al par (N_{max}, ϵ)) se han resuelto con un número de iteraciones menor o igual a max iteraciones. La fila *tiempo total CPU* indica el tiempo total en segundos requerido para calcular φ. La fila *φ obtenido* muestra el mínimo valor de φ obtenido para el cual se ha obtenido una solución factible.

La tabla 2 muestra los resultados para $\epsilon = 0{,}001$ y diferentes valores de N_{max}. Comparando con la tabla 1, se aprecia que para valores pequeños de N_{max}, el número máximo de iteraciones y el tiempo de CPU se degrada claramente conforme ϵ decrece. Los mejores tiempos se obtienen con $N_{max} = 10$ (varios órdenes de magnitud inferiores a los obtenidos con $N_{max} = 1$). Es importante destacar que el valor de φ obtenido para $\epsilon = 0{,}001$ es significativamente menor que el correspondiente a $\epsilon = 0{,}01$.

Finalmente, en la tabla 3, se muestran los resultados para $N_{max} = 10$ y diferentes valores de ϵ. Nótese que en la última fila de esta tabla se muestra el número de iteraciones del método de la bisección que fueron necesarias en cada caso. En contraste con el caso en el que N_{max} es muy pequeño, el comportamiento del algoritmo no se degrada con valores decrecientes de ϵ. Puede observarse que el valor obtenido de φ no mejora significativamente más allá de $\epsilon = 10^{-5}$. Esto significa que la solución óptima se obtiene con $\epsilon = 10^{-5}$ y $N_{max} = 10$ en menos de dos minutos.

Tabla 1: Resultados numéricos para diferentes valores de N_{max} con $\epsilon = 0{,}01$. En estos ejemplos el número de iteraciones del método de la bisección osciló entre 7 y 8.

N_{max}	1	2	5	10
max iteraciones	255	16	8	9
tiempo total CPU	533.4	67.9	21.7	36.2
φ obtenido	2.04	2.09	2.03	2.01

Tabla 2: Resultados para diferentes valores de N_{max} con $\epsilon = 0{,}001$. En estos ejemplos el número de iteraciones del método de la bisección osciló entre 11 y 13.

N_{max}	1	2	5	10
max iteraciones	6021	2726	405	9
tiempo total CPU	$2.4 \cdot 10^4$	$1.2 \cdot 10^4$	$2.2 \cdot 10^3$	63.5
φ obtenido	1.64	1.60	1.59	1.59

Tabla 3: Resultados numéricos para diferentes valores de ϵ con $N_{max} = 10$.

ϵ	10^{-2}	10^{-3}	10^{-4}	10^{-5}	10^{-6}
max iteraciones	9	9	12	9	16
tiempo total CPU	36.2	63.5	103.0	102.3	128.1
φ obtenido	2.01	1.59	1.56	1.55	1.55
iteraciones bisección	7	12	15	19	22

8. Conclusiones

En este trabajo se ha presentado un nuevo algoritmo aleatorio que resuelve un problema de factibilidad robusta con LMIs inciertas. El algoritmo propuesto presenta claras diferencias con otros algoritmos aleatorios gradenciales y de localización. Se garantiza que el algoritmo encuentra una solución factible de nivel δ si el problema es ϵ-factible. Además, si el problema no es ϵ-factible, el algoritmo detecta esta condición en un número finito de iteraciones. Se ha obtenido una cota del número máximo de iteraciones requeridas. Finalmente, los méritos del algoritmo se han ilustrado con un ejemplo numérico.

English Summary

A sequentially optimal randomized algorithm for robust feasibility problems

Abstract This paper proposes a randomized algorithm for feasibility of uncertain LMIs. The algorithm is based on the solution of a sequence of semidefinite optimization problems involving a reduced number of constraints. A bound of the maximum number of iterations required by the algorithm is given. Finally, the performance and behaviour of the algorithm are illustrated by means of a numerical example.
Keywords: robust feasibility linear matrix inequalities randomized algorithms robust control.

Agradecimientos

Los autores agradecen la financiación del Ministerio de Ciencia e Innovación mediante el proyecto DPI2010-21589-C05-01.

Referencias

Alamo, T., Tempo, R., Camacho, E., 2009. Randomized strategies for probabilistic solutions of uncertain feasibility and optimization problems. Automatic Control, IEEE Transactions on 54 (11), 2545 –2559.

Alamo, T., Tempo, R., Ramírez, D., Camacho, E., June 2008. A new vertex result for robustness problems with interval matrix uncertainty. Systems and Control Letters 57 (6), 474–481.

Alamo, T., Tempo, R., Ramirez, D. R., Camacho, E. F., 2007. A sequentially optimal randomized algorithm for robust lmi feasibility problems. In: Proceedings of the European Control Conference. Kos, Greece.

Apkarian, P., Gahinet, P., Becker, G., 1995. Self-scheduled H_∞ control of linear parameter-varying systems: a design example. Automatica 31 (9), 1251 – 1261.
DOI: 10.1016/0005-1098(95)00038-X

Apkarian, P., Tuan, H. D., 2000. Parameterized lmis in control theory. SIAM Journal on Control and Optimization 38 (4), 1241–1264.
DOI: 10.1137/S036301299732612X

Barmish, B., 1994. New Tools for Robustness of Linear Systems. MacMillan Publishing Company, New York, USA.

Barmish, B., Scherbakov, P., December 2000. On avoiding vertexization of robustness problems: The approximate feasibility concept. In: Proceedings of the 39th IEEE Conference on Decision and Control. Sydney, Australia, pp. 1031–1036.

Becker, G., Packard, A., 1994. Robust performance of linear parametrically varying systems using parametrically-dependent linear feedback. Systems and Control Letters 23 (3), 205 – 215.
DOI: 10.1016/0167-6911(94)90006-X

Ben-Tal, A., Nemirovski, A., 2001. Lectures on Modern Convex Optimization. Analysis, Algorithms, and Engineering Applications. MPS/SIAM Series on Optimization, Philadelphia, PA.

Boyd, S., Ghaoui, L. E., Feron, E., Balakrishnan, V., 1994. Linear Matrix Inequalities in Systems and Control Theory. SIAM, Philadelphia, PA.

Calafiore, G., Campi, M., May 2006. The scenario approach to robust control design. IEEE Transactions on Automatic Control 51 (5), 742–753.

Calafiore, G., Dabbene, F., 2007. A probabilistic analytic center cutting plane method for feasibility of uncertain LMIs. Automatica 43, 2022–2033.

Calafiore, G., Dabbene, F., Tempo, R., 2007. A survey of randomized algorithms for control synthesis and performance verification. Journal of Complexity 23 (3), 301 – 316.

Calafiore, G., Polyak, B., November 2001. Stochastic algorithms for exact and approximate feasibility of robust LMIs. IEEE Transactions on Automatic Control 46 (11), 1755–1759.

Calafiore, G. C., Dabbene, F., Tempo, R., 2011. Research on probabilistic methods for control system design. Automatica 47 (7), 1279 – 1293.

Feron, E., Apkarian, P., Gahinet, P., jul 1996. Analysis and synthesis of robust control systems via parameter-dependent lyapunov functions. Automatic Control, IEEE Transactions on 41 (7), 1041 –1046.

Fujisaki, Y., Dabbene, F., Tempo, R., 2003. Probabilistic design of lpv control systems. Automatica 39, 1323–1337.

Kanev, S., Schutter, B. D., Verhaegen, M., 2003. An ellipsoid algorithm for probabilistic robust controller design. Systems and Control Letters 49, 365–375.

Liberzon, D., Tempo, R., June 2004. Common Lyapunov functions and gradient algorithms. IEEE Transactions on Automatic Control 49 (6), 990–994.

Nemirovskii, A., 1993. Several NP-hard problems arising in robust stability analysis. Mathematics of Control, Signal and Systems 6 (2), 99–105.

Oishi, Y., December 2003. Probabilistic design of a robust state-feedback controller based on parameter-dependent Lyapunov functions. In: Proceedings of the 42nd IEEE Conference on Decision and Control. Maui, Hawaii USA, pp. 1920–1925.

Polyak, B., Tempo, R., 2001. Probabilistic robust design with linear quadratic regulators. Systems and Control Letters 43 (5), 343–353.

Scokaert, P. O. M., Mayne, D. Q., Rawlings, J. B., 1999. Suboptimal model predictive control (feasibility implies stability). IEEE Transactions on Automatic Control 44 (3), 648–654.

Tempo, R., Calafiore, G., Dabbene, F., 2005. Randomized Algorithms for Analysis and Control of Uncertain Systems. Communications and Control Engineering Series. Springer-Verlag, London.

Tuan, H., Apkarian, P., Nguyen, T., dec 2001. Robust and reduced-order filtering: new lmi-based characterizations and methods. Signal Processing, IEEE Transactions on 49 (12), 2975 –2984.
DOI: 10.1109/78.969506

Wan, Z., Kothare, M., 2002. Robust output feedback model predictive control using off-line linear matrix inequalities. Journal of Process Control 12, 763–774.

Zhou, K., Doyle, J., Glover, K., 1996. Robust and Optimal Control. Prentice-Hall, Englewood Cliff, NJ.

Apéndice A. Propiedades auxiliares

Propiedad 5. *Supóngase que* $\mathbf{h} \in \mathrm{Co}\{\mathcal{C}\}$ *y que* $W_c \subset W$. *Considérese ahora el siguiente problema de optimización*

$$\begin{array}{rcl} \gamma_c^* = & \min\limits_{\mathbf{z},\gamma} & \gamma \\ & s.a. & \mathbf{h}^\top \mathbf{z} \le \gamma \\ & & A(\mathbf{z}, w) \le \gamma \mathrm{I}, \quad \forall w \in W_c \\ & & \mathbf{z}^\top \mathbf{z} \le 1. \end{array} \tag{A.1}$$

Denótese por $(\gamma_c^*, \mathbf{z}_c^*)$ *la solución óptima del problema anterior. Entonces,*

(I) $\gamma_c^* z_c^* \in \mathrm{Co}\{\mathcal{C}\}$
(II) $\gamma_c^* \le \gamma^*$
(III) *Si* $\gamma_c^* \neq 0$ *entonces* $\|z_c^*\|_2 = 1$.

Prueba: Teniendo en cuenta la propiedad 1 es posible concluir que, si $(\mathbf{z}_c^*, \gamma_c^*)$ es la solución del problema (A.1), entonces también lo es de

$$\begin{array}{rcl} \gamma_c^* = & \min\limits_{\mathbf{z},\gamma} & \gamma \\ & s.a. & \mathbf{h}^\top \mathbf{z} \le \gamma \\ & & \mathbf{c}^\top \mathbf{z} \le \gamma, \quad \forall \mathbf{c} \in \bigcup\limits_{w \in W_c} C(w) \\ & & \mathbf{z}^\top \mathbf{z} \le 1. \end{array} \tag{A.2}$$

Denótese $T_c = \mathbf{h} \bigcup \left(\bigcup\limits_{w \in W_c} C(w) \right)$. Como $C(w) \subseteq \mathcal{C} \subseteq \mathrm{Co}\{\mathcal{C}\}$, resulta que $T_c \subseteq \mathrm{Co}\{\mathcal{C}\}$. Es decir, el problema de optimización (A.2) puede reescribirse como:

$$\begin{array}{rcl} \gamma_c^* = & \min\limits_{\mathbf{z},\gamma} & \gamma \\ & s.a. & \mathbf{c}^\top \mathbf{z} \le \gamma, \quad \forall \mathbf{c} \in T_c \\ & & \mathbf{z}^\top \mathbf{z} \le 1. \end{array} \tag{A.3}$$

En general, T_c tiene un número infinito de elementos, lo que implica que se deberían considerar un número infinito de restricciones en la resolución del problema. Sin embargo, sólo un número finito de restricciones condiciona la solución del problema (Calafiore and Campi (2006)), es decir, existirá un número finito de restricciones

$$\mathbf{c}_i^\top \mathbf{z} \le \gamma,\ i = 1, \ldots, r$$

con $\mathbf{c}_i \in T_c \subseteq \mathrm{Co}\{C\}$, $i = 1, \ldots, r$ que determinarán la solución del problema (A.3). Esto significa que $(\mathbf{z}_c^*, \gamma_c^*)$ se obtienen del siguiente problema de optimización

$$\begin{aligned} \gamma_c^* = \ & \min_{\mathbf{z},\gamma} \ \gamma \\ & s.a. \ \ \mathbf{c}_i^\top \mathbf{z} \le \gamma, \ \ i = 1, \ldots, r \\ & \qquad \mathbf{z}^\top \mathbf{z} \le 1. \end{aligned} \tag{A.4}$$

Teniendo en cuenta la propiedad 8 se infiere que

$$\mathbf{z}_c^* = \begin{cases} \dfrac{-\tilde{\mathbf{c}}}{\sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}}} & \text{if} \ \ \tilde{\mathbf{c}} \neq 0 \\ 0 & \text{if} \ \ \tilde{\mathbf{c}} = 0 \end{cases}, \quad \gamma_c^* = -\|\tilde{\mathbf{c}}\|_2 \tag{A.5}$$

donde $\tilde{\mathbf{c}}$ es el vector de $\Gamma = \mathrm{Co}\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_r\} \subseteq \mathrm{Co}\{C\}$ con menor norma euclídea.

Supóngase que $\tilde{\mathbf{c}} = 0$. Esto significa que $0 \in \Gamma$. En ese caso, $\gamma_c^* \mathbf{z}_c^* = 0 \in \Gamma \subseteq \mathrm{Co}\{C\}$. Supóngase ahora que $\tilde{\mathbf{c}} \neq 0$, entonces

$$\gamma_c^* \mathbf{z}_c^* = (-\sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}}) \left(\frac{-\tilde{\mathbf{c}}}{\sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}}} \right) = \tilde{\mathbf{c}} \in \Gamma \subseteq \mathrm{Co}\{C\}.$$

De esta manera se ha probado la primera parte de esta propiedad.

Nótese que $\tilde{\mathbf{c}} \in \mathrm{Co}\{C\}$ y $\gamma_c^* = -\|\tilde{\mathbf{c}}\|_2$. De esto y de la propiedad 3 se infiere que $\gamma^* \ge \gamma_c^*$. Esto prueba la segunda parte de esta propiedad.

Finalmente, nótese que $\gamma_c^* = -\|\tilde{\mathbf{c}}\|_2 \neq 0$ implica que $\tilde{\mathbf{c}} \neq 0$. En ese caso, se deriva de la ecuación (A.5) que:

$$\|\mathbf{z}_c^*\|_2 = \left\| \frac{-\tilde{\mathbf{c}}}{\sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}}} \right\|_2 = 1. \qquad \square$$

Propiedad 6. *Supóngase que $W_c \subset W$. Considérese ahora el siguiente problema de optimización*

$$\begin{aligned} \gamma_c^* = \ & \min_{\mathbf{z},\gamma} \ \gamma \\ & s.a. \ \ A(\mathbf{z}, w) \le \gamma \mathrm{I}, \ \ \forall w \in W_c \\ & \qquad \mathbf{z}^\top \mathbf{z} \le 1. \end{aligned}$$

Denótese como $(\gamma_c^*, \mathbf{z}_c^*)$ la solución óptima de este problema. Entonces,

(i) $\gamma_c^* z_c^* \in \mathrm{Co}\{C\}$

(ii) $\gamma_c^* \le \gamma^*$

(iii) *Si* $\gamma_c^* \neq 0$ *entonces* $\|z_c^*\|_2 = 1$.

Prueba: Al igual que en la demostración de la propiedad 5, se puede probar que $(\gamma_c^*, \mathbf{z}_c^*)$ es la solución del problema de optimización

$$\begin{aligned} \gamma_c^* = \ & \min_{\mathbf{z},\gamma} \ \gamma \\ & s.a. \ \ \mathbf{c}^\top \mathbf{z} \le \gamma, \ \ \forall \mathbf{c} \in T_c \\ & \qquad \mathbf{z}^\top \mathbf{z} \le 1 \end{aligned}$$

donde $T_c = \bigcup_{w \in W_c} C(w) \subseteq \mathrm{Co}\{C\}$. El resto de la demostración es idéntica a la de la propiedad 5 y por tanto se omite aquí. $\square$

Propiedad 7. *Supóngase que un problema de factibilidad robusta (1) es ϵ-factible. Supóngase también que $\max_{\mathbf{c} \in C} \|\mathbf{c}\|_2 < \sigma$. Entonces, existe $\bar{\mathbf{z}}$, $\bar{\mathbf{z}}^\top \bar{\mathbf{z}} \le 1$ tal que*

$$A(\mathbf{z}, w) \le \frac{-\epsilon}{2} \mathrm{I}, \quad \forall w \in W, \ \forall \mathbf{z} \in \{ \ \mathbf{z} \ : \ \|\mathbf{z} - \bar{\mathbf{z}}\|_2 \le \frac{\epsilon}{2\sigma} \ \}.$$

Prueba: Como se asume que el problema es ϵ factible, entonces existe $\bar{\mathbf{z}}$ tal que $\bar{\mathbf{z}}^\top \bar{\mathbf{z}} \le 1$ y $\bar{\lambda}(A(\bar{\mathbf{z}}, w)) \le -\epsilon$, $\forall w \in W$. Supóngase ahora que $\mathbf{z} \in \{ \ \mathbf{z} \ : \ \|\mathbf{z} - \bar{\mathbf{z}}\|_2 \le \frac{\epsilon}{2\sigma} \ \}$ y que $w \in W$. La propiedad (1) garantiza que

$$\begin{aligned} \bar{\lambda}(A(\mathbf{z}, w)) &= \max_{\mathbf{c} \in C(w)} \mathbf{c}^\top \mathbf{z} \\ &= \max_{\mathbf{c} \in C(w)} \mathbf{c}^\top (\bar{\mathbf{z}} + (\mathbf{z} - \bar{\mathbf{z}})) \\ &\le \max_{\mathbf{c} \in C(w)} \mathbf{c}^\top \bar{\mathbf{z}} + \max_{\mathbf{c} \in C(w)} \mathbf{c}^\top (\mathbf{z} - \bar{\mathbf{z}}) \\ &\le \bar{\lambda}(A(\bar{\mathbf{z}}, w)) + \max_{\mathbf{c} \in C(w)} \|\mathbf{c}\|_2 \|\mathbf{z} - \bar{\mathbf{z}}\|_2 \\ &\le -\epsilon + \max_{\mathbf{c} \in C} \|\mathbf{c}\|_2 \|\mathbf{z} - \bar{\mathbf{z}}\|_2 \\ &\le -\epsilon + \sigma \frac{\epsilon}{2\sigma} = -\frac{\epsilon}{2}. \end{aligned} \qquad \square$$

Propiedad 8. *Considérese el problema de optimización:*

$$\begin{aligned} \gamma_c^* = \ & \min_{\mathbf{z},\gamma} \ \gamma \\ & s.a. \ \ \mathbf{c}_i^\top \mathbf{z} \le \gamma, \ \ i = 1, \ldots, r \\ & \qquad \mathbf{z}^\top \mathbf{z} \le 1. \end{aligned} \tag{A.6}$$

Denótese por $(\mathbf{z}_c^*, \gamma_c^*)$ la solución de (A.6). Denótese también como $\tilde{\mathbf{c}}$ el vector de $\Gamma = \mathrm{Co}\{\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_r\}$ con la menor norma euclídea. Entonces,

(i) $\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}} - \mathbf{c}_i^\top \tilde{\mathbf{c}} \le 0, \ i = 1, \ldots, r$

(ii)

$$\mathbf{z}_c^* = \begin{cases} \dfrac{-\tilde{\mathbf{c}}}{\sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}}} & \textit{if} \ \ \tilde{\mathbf{c}} \neq 0 \\ 0 & \textit{if} \ \ \tilde{\mathbf{c}} = 0 \end{cases}, \quad \gamma_c^* = -\|\tilde{\mathbf{c}}\|_2.$$

Prueba:

(i) Supóngase que existe un valor j tal que $1 \le j \le r$ y $\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}} - \mathbf{c}_j^\top \tilde{\mathbf{c}} > 0$. Se demostrará que esto contradice la suposición de que $\tilde{\mathbf{c}}$ es el vector de norma euclídea mínima en Γ. Denótese como $J(\alpha) = \|(1 - \alpha)\tilde{\mathbf{c}} + \alpha \mathbf{c}_j\|_2^2$. Entonces

$$\frac{d}{d\alpha} J(\alpha) = \frac{d}{d\alpha} \|\tilde{\mathbf{c}} - \alpha(\tilde{\mathbf{c}} - \mathbf{c}_j)\|_2^2 =$$
$$2\alpha \|\tilde{\mathbf{c}} - \mathbf{c}_j\|_2^2 - 2(\tilde{\mathbf{c}} - \mathbf{c}_j)^\top \tilde{\mathbf{c}}.$$

Por tanto, $\frac{d}{d\alpha} J(\alpha)\big|_{\alpha=0} = -2(\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}} - \mathbf{c}_j^\top \tilde{\mathbf{c}}) < 0$. Esto significa que existe un $\alpha^* \in (0, 1]$ tal que $\|\tilde{\mathbf{c}}\|_2^2 = J(0) > J(\alpha^*) = \|(1 - \alpha^*)\tilde{\mathbf{c}} + \alpha^* \mathbf{c}_j\|_2^2$. Como $(1 - \alpha^*)\tilde{\mathbf{c}} + \alpha^* \mathbf{c}_j \in \Gamma$, la desigualdad $J(0) > J(\alpha^*)$ contradice la suposición de que $\tilde{\mathbf{c}}$ es el vector de Γ con menor norma euclidea.

(ii) Como $\tilde{\mathbf{c}} \in \Gamma$, entonces existen (por definición de envoltura convexa) $\lambda_1, \lambda_2, \ldots, \lambda_r$ tales que $\lambda_i \ge 0$, $i = 1, \ldots, r$, $\sum_{i=1}^{r} \lambda_i \mathbf{c}_i = \tilde{\mathbf{c}}$ y $\sum_{i=1}^{r} \lambda_i = 1$. Nótese que $\mathbf{c}_i^\top \mathbf{z} \le \gamma$, $i = 1, \ldots, r$

implica que $\sum_{i=1}^{r} \lambda_i \mathbf{c}_i^\top \mathbf{z} = \tilde{\mathbf{c}}^\top \mathbf{z} \leq \sum_{i=1}^{r} \lambda_i \gamma = \gamma$. De esto se infiere que

$$\begin{array}{rll} \gamma_c^* \geq & \min\limits_{\mathbf{z},\gamma} & \gamma \\ & s.a. & \tilde{\mathbf{c}}^\top \mathbf{z} \leq \gamma \\ & & \mathbf{z}^\top \mathbf{z} \leq 1. \end{array} \tag{A.7}$$

En este punto se han de considerar dos casos:

- $\tilde{\mathbf{c}} = 0$. En este caso el problema de optimización (A.7) se reescribe como:

$$\begin{array}{rll} \gamma_c^* \geq & \min\limits_{\mathbf{z},\gamma} & \gamma \\ & s.a. & 0 \leq \gamma \\ & & \mathbf{z}^\top \mathbf{z} \leq 1. \end{array} \tag{A.8}$$

  De lo anterior se deduce que $\gamma_c^* \geq 0$. Nótese que $(\mathbf{z}, \gamma) = (0, 0)$ es una solución factible al problema de optimización (A.6), lo que implica que $\gamma_c^* \leq 0$. De las desigualdades obtenidas: $\gamma_c^* \geq 0$, $\gamma_c^* \leq 0$ se concluye que $\gamma_c^* = 0$. Esto significa que la solución del problema (A.6) se da en $(\mathbf{z}_c^*, \gamma_c^*) = (0, 0) = (0, -\|\tilde{\mathbf{c}}\|_2)$.

- $\tilde{\mathbf{c}} \neq 0$ Nótese que $\tilde{\mathbf{z}} = \dfrac{-\tilde{\mathbf{c}}}{\sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}}}$ es el vector que minimiza $\tilde{\mathbf{c}}^\top \mathbf{z}$ sujeto a la restricción $\mathbf{z}^\top \mathbf{z} \leq 1$. Esto significa que la solución al problema de optimización (A.7) se alcanza en $(\tilde{\mathbf{z}}^*, \tilde{\gamma}^*) = \left(\dfrac{-\tilde{\mathbf{c}}}{\sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}}}, -\sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}} \right)$. Por tanto,

$$\gamma_c^* \geq -\sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}}. \tag{A.9}$$

  Además para todo i, $1 \leq i \leq r$

$$\begin{aligned} \mathbf{c}_i^\top \tilde{\mathbf{z}}^* + \sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}} &= \frac{-\mathbf{c}_i^\top \tilde{\mathbf{c}}}{\sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}}} + \sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}} \\ &= \frac{1}{\sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}}} \left(\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}} - \mathbf{c}_i^\top \tilde{\mathbf{c}} \right) \leq 0. \end{aligned}$$

  La última desigualdad se debe a la primera parte de la propiedad que se está demostrando. Por tanto se ha demostrado que

$$\mathbf{c}_i^\top \tilde{\mathbf{z}}^* \leq -\sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}}, \quad i = 1, \ldots, r.$$

  Esto, junto que el hecho de que $(\tilde{\mathbf{z}}^*)^\top \tilde{\mathbf{z}}^* = 1$ implica que $(\tilde{\mathbf{z}}^*, -\sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}})$ es una solución factible de (A.6). Teniendo en cuenta la ecuación (A.9), se infiere que $(\tilde{\mathbf{z}}^*, -\sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}}) = (\frac{-\tilde{\mathbf{c}}}{\sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}}}, -\sqrt{\tilde{\mathbf{c}}^\top \tilde{\mathbf{c}}})$ es la solución factible óptima del problema de optimización (A.6). □

Propiedad 9. *Sea* $\mathbf{h} \in \mathrm{Co}\{C\}/0$ *y* $\mathbf{c} \in C$. *Supóngase entonces que* $\mathbf{c}^\top \mathbf{h} < 0$ *y que* $\max\limits_{\mathbf{c} \in C} \|\mathbf{c}\|_2 < \sigma$. *Entonces,*

1. *La solución del problema de optimización*

$$\min_{\lambda \in [0,1]} \|(1-\lambda)\mathbf{h} + \lambda \mathbf{c}\|_2$$

   se da en $\lambda^* = \dfrac{(\mathbf{h}-\mathbf{c})^\top \mathbf{h}}{\|\mathbf{h}-\mathbf{c}\|_2^2} \in [0, 1]$

2. $\|(1-\lambda^*)\mathbf{h} + \lambda^* \mathbf{c}\|_2 \leq \|\mathbf{h}\|_2 \left(1 - \frac{\|\mathbf{h}\|_2^2}{8\sigma^2}\right)$

Prueba:

1. De $\mathbf{c}^\top \mathbf{h} < 0$, se infiere directamente que:
   a) $\mathbf{h} \neq \mathbf{c}$.
   b) $(\mathbf{h}-\mathbf{c})^\top \mathbf{h} > 0$.
   c) $(\mathbf{h}-\mathbf{c})^\top \mathbf{h} < (\mathbf{h}-\mathbf{c})^\top \mathbf{h} - \mathbf{c}^\top \mathbf{h} + \mathbf{c}^\top \mathbf{c} = \|\mathbf{h}-\mathbf{c}\|_2^2$.

   De estos tres puntos resulta que

$$\lambda_d = \frac{(\mathbf{h}-\mathbf{c})^\top \mathbf{h}}{\|\mathbf{h}-\mathbf{c}\|_2^2} \in [0, 1]$$

   Nótese que el valor óptimo de λ minimiza

$$\begin{gathered} J(\lambda) = ((1-\lambda)\mathbf{h} + \lambda \mathbf{c})^\top ((1-\lambda)\mathbf{h} + \lambda \mathbf{c}) = \\ (1-\lambda)^2 \mathbf{h}^\top \mathbf{h} + 2\lambda(1-\lambda)\mathbf{c}^\top \mathbf{h} + \lambda^2 \mathbf{c}^\top \mathbf{c} = \\ \lambda^2(\mathbf{h}^\top \mathbf{h} - 2\mathbf{c}^\top \mathbf{h} + \mathbf{c}^\top \mathbf{c}) + 2\lambda(\mathbf{c}^\top \mathbf{h} - \mathbf{h}^\top \mathbf{h}) + \mathbf{h}^\top \mathbf{h} = \\ \lambda^2 \|\mathbf{h}-\mathbf{c}\|_2^2 + 2\lambda(\mathbf{c}-\mathbf{h})^\top \mathbf{h} + \mathbf{h}^\top \mathbf{h} \end{gathered}$$

   Nótese también, que la expresión anterior es convexa en λ. La derivada con respecto a λ es

$$\frac{d}{d\lambda} J(\lambda) = 2\lambda \|\mathbf{h}-\mathbf{c}\|_2^2 + 2(\mathbf{c}-\mathbf{h})^\top \mathbf{h} \tag{A.10}$$

   Por tanto, la derivada se anula en $\lambda_d = \dfrac{(\mathbf{h}-\mathbf{c})^\top \mathbf{h}}{\|\mathbf{h}-\mathbf{c}\|_2^2}$. Tal y como se ha mostrado anteriormente, $\lambda_d \in [0, 1]$. De esto y de la convexidad con respecto a λ se concluye que el mínimo (sujeto a la restricción $\lambda \in [0, 1]$) se da en $\lambda^* = \lambda_d$.

2. Denótese por $\tilde{\mathbf{h}} = (1-\lambda^*)\mathbf{h} + \lambda^* \mathbf{c}$. Téngase en cuenta que $\|\tilde{\mathbf{h}}\|_2^2 = J(\lambda^*)$. Por tanto,

$$\begin{aligned} \|\tilde{\mathbf{h}}\|_2^2 &= (\lambda^*)^2 \|\mathbf{h}-\mathbf{c}\|_2^2 + 2\lambda^*(\mathbf{c}-\mathbf{h})^\top \mathbf{h} + \mathbf{h}^\top \mathbf{h} \\ &= \frac{|(\mathbf{h}-\mathbf{c})^\top \mathbf{h}|^2}{\|\mathbf{h}-\mathbf{c}\|_2^2} - 2\frac{|(\mathbf{h}-\mathbf{c})^\top \mathbf{h}|^2}{\|\mathbf{h}-\mathbf{c}\|_2^2} + \mathbf{h}^\top \mathbf{h} \\ &= \|\mathbf{h}\|_2^2 - \frac{|(\mathbf{h}-\mathbf{c})^\top \mathbf{h}|^2}{\|\mathbf{h}-\mathbf{c}\|_2^2}. \end{aligned}$$

   De ahí podemos obtener

$$\begin{aligned} \|\mathbf{h}\|_2 - \|\tilde{\mathbf{h}}\|_2 &= \sqrt{\mathbf{h}^\top \mathbf{h}} - \sqrt{\mathbf{h}^\top \mathbf{h} - \frac{|(\mathbf{h}-\mathbf{c})^\top \mathbf{h}|^2}{\|\mathbf{h}-\mathbf{c}\|_2^2}} \\ &= \frac{\frac{|(\mathbf{h}-\mathbf{c})^\top \mathbf{h}|^2}{\|\mathbf{h}-\mathbf{c}\|_2^2}}{\sqrt{\mathbf{h}^\top \mathbf{h}} + \sqrt{\mathbf{h}^\top \mathbf{h} - \frac{|(\mathbf{h}-\mathbf{c})^\top \mathbf{h}|^2}{\|\mathbf{h}-\mathbf{c}\|_2^2}}} \\ &> \frac{|(\mathbf{h}-\mathbf{c})^\top \mathbf{h}|^2}{2\sqrt{\mathbf{h}^\top \mathbf{h}}\|\mathbf{h}-\mathbf{c}\|_2^2} \\ &> \frac{|\mathbf{h}^\top \mathbf{h}|^2}{2\sqrt{\mathbf{h}^\top \mathbf{h}}\|\mathbf{h}-\mathbf{c}\|_2^2} \\ &\geq \frac{\|\mathbf{h}\|_2^3}{2(\|\mathbf{h}\|_2 + \|\mathbf{c}\|_2)^2} \\ &\geq \frac{\|\mathbf{h}\|_2^3}{2(2\sigma)^2} = \frac{\|\mathbf{h}\|_2^3}{8\sigma^2}. \end{aligned}$$

Esto prueba que

$$\|\mathbf{h}\|_2 - \|\tilde{\mathbf{h}}\|_2 \geq \frac{\|\mathbf{h}\|_2^3}{8\sigma^2} \quad \text{y que} \quad \|\tilde{\mathbf{h}}\|_2 \leq \|\mathbf{h}\|_2\left(1 - \frac{\|\mathbf{h}\|_2^2}{8\sigma^2}\right) \quad \square$$

Propiedad 10. *Considérese la secuencia $\Gamma_0, \Gamma_1, \ldots, \Gamma_k, \ldots,$ y supóngase que*

- $\Gamma_0 < \sigma$,
- $0 \leq \Gamma_{k+1} \leq \Gamma_k(1 - \tau\Gamma_k^2)$, $\forall k \geq 0$
- $\sigma > \epsilon > 0$ *y* $0 < \tau\epsilon^2 < 1$

Entonces,

- $\Gamma_k < \epsilon$ *para todo* $k \geq \left\lceil \dfrac{\ln \frac{\sigma}{\epsilon}}{\tau\epsilon^2} \right\rceil$ *y* $k \geq \left\lceil \frac{1}{2\tau\epsilon^2} \right\rceil + \left\lceil \frac{\ln \ln \frac{\sigma^2}{\epsilon^2}}{2\tau\epsilon^2} \right\rceil$.

Prueba:

- $\Gamma_k < \epsilon$ para cada $k \geq \left\lceil \dfrac{\ln \frac{\sigma}{\epsilon}}{\tau\epsilon^2} \right\rceil$:

  Supóngase que $\Gamma_j \geq \epsilon$, $j = 0, \ldots, k-1$. Entonces

  $$\Gamma_{j+1} \leq \Gamma_j(1 - \tau\Gamma_j^2) \leq \Gamma_j(1 - \tau\epsilon^2), \quad j = 0, \ldots, k-1$$

  De ahí,

  $$\Gamma_k \leq \Gamma_0(1 - \tau\epsilon^2)^k < \sigma(1 - \tau\epsilon^2)^k.$$

  Por tanto, la desigualdad $\Gamma_k < \epsilon$ se cumple para cada $k \geq k^*$, si k^* satisface

  $$\sigma(1 - \tau\epsilon^2)^{k^*} \leq \epsilon$$

  o de manera equivalente

  $$k^* \ln(1 - \tau\epsilon^2) \leq \ln \frac{\epsilon}{\sigma}.$$

  Teniendo en cuenta que $\ln(1 - \tau\epsilon^2) < 0$,

  $$k^* \geq \frac{\ln \frac{\epsilon}{\sigma}}{\ln(1 - \tau\epsilon^2)} = \frac{\ln \frac{\sigma}{\epsilon}}{-\ln(1 - \tau\epsilon^2)}.$$

  Nótese que $-\ln(1-x)$ es una función convexa de x en el dominio $x < 1$. Esto implica que $-\ln(1-x)$ es mayor o igual que la linealización de $-\ln(1-x)$ en $x = 0$. Esto es, $-\ln(1-x) \geq \ln(1-0) + \frac{d}{dx}(-\ln(1-x))\big|_{x=0}(x-0) = x$. Por tanto, $-\ln(1 - \tau\epsilon^2) \geq \tau\epsilon^2$. Esto significa que $\Gamma_k < \epsilon$ para todo k que satisfaga

  $$k \geq \left\lceil \frac{\ln \frac{\sigma}{\epsilon}}{\tau\epsilon^2} \right\rceil \geq \left\lceil \frac{\ln \frac{\sigma}{\epsilon}}{-\ln(1 - \tau\epsilon^2)} \right\rceil.$$

- $\Gamma_k < \epsilon$ para todo $k \geq \left\lceil \frac{1}{2\tau\epsilon^2} \right\rceil + \left\lceil \frac{\ln \ln \frac{\sigma^2}{\epsilon^2}}{2\tau\epsilon^2} \right\rceil$:

  En primer lugar, si $\ln \frac{\sigma}{\epsilon} \leq \frac{1}{2}$ entonces la aplicación directa del primer resultado de esta propiedad lleva a: $\Gamma_k < \epsilon$, para todo $k \geq \left\lceil \dfrac{1}{2\tau\epsilon^2} \right\rceil$. Es decir, queda por probar que se cumple esta parte de la propiedad bajo la suposición de que $\ln \frac{\sigma}{\epsilon} > \frac{1}{2}$.

  Supóngase que c satisface: $\sigma > c > \epsilon$. Denótese $k_{ac} = \left\lceil \dfrac{\ln \frac{\sigma}{c}}{\tau c^2} \right\rceil$. Usando la primera parte de esta propiedad se infiere que $\Gamma_k < c$ para cada $k \geq k_{ac}$. Denótese $k_{cb} = \left\lceil \dfrac{\ln \frac{c}{\epsilon}}{\tau\epsilon^2} \right\rceil$, entonces, teniendo en cuenta que $\Gamma_{k_{ac}} < c$ y usando de nuevo la primera parte de la propiedad: $\Gamma_{k_{ac}+j} < \epsilon$ para todo $j \geq k_{bc}$. De igual manera, $\Gamma_k < \epsilon$ para todo

  $$k \geq k_{ac} + k_{bc} = \left\lceil \frac{\ln \frac{\sigma}{c}}{\tau c^2} \right\rceil + \left\lceil \frac{\ln \frac{c}{\epsilon}}{\tau\epsilon^2} \right\rceil.$$

  Supóngase ahora que $c = \epsilon\sqrt{\ln \dfrac{\sigma^2}{\epsilon^2}}$. Se demostrará a continuación que este valor de c satisface: $\sigma > c > \epsilon$. Como $\ln \frac{\sigma}{\epsilon} > \frac{1}{2}$ resulta que $\ln \frac{\sigma^2}{\epsilon^2} > 1$. Por tanto, $c = \epsilon\sqrt{\ln \dfrac{\sigma^2}{\epsilon^2}} > \epsilon$. Por otra parte, la desigualdad $\ln x < x$, $\forall x > 1$ implica que $\ln x^2 < x^2$, $\forall x > 1$. De igual modo, $\sqrt{\ln x^2} < x$. De esto se tiene que: $\sqrt{\ln \frac{\sigma^2}{\epsilon^2}} < \frac{\sigma}{\epsilon}$, lo que implica que: $c = \epsilon\sqrt{\ln \dfrac{\sigma^2}{\epsilon^2}} < \sigma$.

  Ahora se obtendrán cotas de los valores de k_{ac} y k_{cb} para c. Recuérdese que se ha asumido que $\ln \frac{\sigma}{\epsilon} > \frac{1}{2}$. De ahí, $\sqrt{\ln \dfrac{\sigma^2}{\epsilon^2}} > 1$, y

  $$k_{ac} = \left\lceil \frac{\ln \frac{\sigma}{c}}{\tau c^2} \right\rceil = \left\lceil \frac{\ln\left(\frac{\sigma}{\epsilon\sqrt{\ln \frac{\sigma^2}{\epsilon^2}}}\right)}{\tau\epsilon^2 \ln \frac{\sigma^2}{\epsilon^2}} \right\rceil \leq \left\lceil \frac{\ln \frac{\sigma}{\epsilon}}{2\tau\epsilon^2 \ln \frac{\sigma}{\epsilon}} \right\rceil = \left\lceil \frac{1}{2\tau\epsilon^2} \right\rceil,$$

  $$k_{cb} = \left\lceil \frac{\ln \frac{c}{\epsilon}}{\tau\epsilon^2} \right\rceil = \left\lceil \frac{\ln \sqrt{\ln \frac{\sigma^2}{\epsilon^2}}}{\tau\epsilon^2} \right\rceil = \left\lceil \frac{\ln \ln \frac{\sigma^2}{\epsilon^2}}{2\tau\epsilon^2} \right\rceil.$$

  $\square$

Apéndice B. Demostraciones de algunas propiedades

Apéndice B.1. Prueba del teorema 1

(i) En primer lugar se probará que si $\gamma^* \neq 0$ entonces el problema (1) es factible. Nótese que el problema de optimización (4) es siempre factible ($\mathbf{z} = 0$, $\gamma = 0$ es una solución factible). Dado que $\gamma = 0$ forma parte de una solución factible, el valor óptimo de γ será $\gamma^* \leq 0$. De esto y de la suposición de que $\gamma^* \neq 0$ se infiere que γ^* es estrictamente menor que cero. Esto es, si el mínimo del problema de optimización se da en $(\mathbf{z}^*, \gamma^*)$ entonces $\gamma^* < 0$ y

$$A(\mathbf{z}^*, w) \leq \gamma^* \mathrm{I} < 0, \quad \forall w \in W.$$

De esto se concluye que $\mathbf{z}^*$ es una solución factible del problema (1).

Para finalizar la prueba de la primera parte del teorema es necesario mostrar que $\gamma^* = 0$ implica que el problema (1) es no factible. Este hecho se probará mostrando que el problema de factibilidad (1) implica que $\gamma^* < 0$.

Supóngase que existe $\bar{\mathbf{z}}$ tal que satisface las restricciones robustas de (1). Esto es, $\bar{\mathbf{z}}$ pertenece a $\mathcal{D}$. La desigualdad estricta del problema (1) implica que $\bar{\mathbf{z}} \neq 0$. Teniendo en cuenta la dependencia lineal con respecto a $\mathbf{z}$, resulta que $\rho\bar{\mathbf{z}} \in D$ para todo $\rho > 0$, $\rho \in \mathbb{R}$. En particular, $\hat{\mathbf{z}} = \left(\frac{1}{\sqrt{\bar{\mathbf{z}}^\top\bar{\mathbf{z}}}}\right)\bar{\mathbf{z}} \in \mathcal{D}$. Nótese que $\hat{\mathbf{z}} \in \mathcal{D}$ y $\hat{\mathbf{z}}^\top\hat{\mathbf{z}} = 1$. De esto y de la compacidad de W se infiere que existe un $\hat{\gamma} < 0$ tal que

$$\begin{array}{rcl} A(\hat{\mathbf{z}}, w) & \leq & \hat{\gamma}\mathrm{I}, \quad \forall w \in W \\ \hat{\mathbf{z}}^\top\hat{\mathbf{z}} & \leq & 1. \end{array}$$

Por tanto, el γ óptimo será al menos igual o menor que $\hat{\gamma}$ por lo que $\gamma^* < 0$. □

(ii) Esto se deriva directamente de la definición de ϵ-factibilidad. Supóngase que $\epsilon > -\gamma^*$ y que el problema es ϵ-factible. Entonces, existe $\hat{\mathbf{z}}$ tal que $A(\hat{\mathbf{z}}, w) \leq -\epsilon\mathrm{I}$, $\forall w \in W$ y $\hat{\mathbf{z}}^\top\hat{\mathbf{z}} \leq 1$. Por tanto, $\gamma^* \leq -\epsilon$. Esto contradice la suposición de que $\epsilon > -\gamma^*$. □

Apéndice B.2. Prueba de la propiedad 3

Supóngase que $\bar{\mathbf{h}}$ pertenece a $\mathrm{Co}\{\mathcal{C}\}$. Se demostrará que $\gamma^* \geq -\|\bar{\mathbf{h}}\|_2$. Dado que $\bar{\mathbf{h}} \in \mathrm{Co}\{\mathcal{C}\}$ se infiere que existen $\mathbf{c}_1, \ldots, \mathbf{c}_p$ y $\lambda_1, \ldots, \lambda_p$ tales que: $\bar{\mathbf{h}} = \sum_{i=1}^{p} \lambda_i\mathbf{c}_i$, $1 = \sum_{i=1}^{p} \lambda_i$, $\lambda_i \geq 0$, $\mathbf{c}_i \in \mathcal{C}$, $i = 1, \ldots, p$. De (5) se infiere que

$$\begin{array}{rll} \gamma^* \geq \gamma_c^* = & \min\limits_{\mathbf{z},\gamma} & \gamma \\ & s.a. & \mathbf{c}_i^\top\mathbf{z} \leq \gamma, \quad i = 1, \ldots, p \\ & & \mathbf{z}^\top\mathbf{z} \leq 1, \end{array}$$

debido a que es el mismo problema de minimización que (5) pero con menos restricciones, por lo que el mínimo será menor o igual. Nótese que las desigualdades $\mathbf{c}_i^\top\mathbf{z} \leq \gamma$, $i = 1, \ldots, p$ implican que

$$\sum_{i=1}^{p} \lambda_i\mathbf{c}_i^\top\mathbf{z} \leq \sum_{i=1}^{p} \lambda_i\gamma = \gamma \quad \text{y que} \quad \left(\sum_{i=1}^{p} \lambda_i\mathbf{c}_i\right)^\top \mathbf{z} = \bar{\mathbf{h}}^\top\mathbf{z} \leq \gamma.$$

Por tanto,

$$\begin{array}{rll} \gamma^* \geq \gamma_c^* \geq \gamma_d^* = & \min\limits_{\mathbf{z},\gamma} & \gamma \\ & s.a. & \bar{\mathbf{h}}^\top\mathbf{z} \leq \gamma \\ & & \mathbf{z}^\top\mathbf{z} \leq 1. \end{array}$$

Nótese que la solución al problema de optimización anterior es

$$\mathbf{z}_d^* = \begin{cases} \frac{-\bar{\mathbf{h}}}{\sqrt{\bar{\mathbf{h}}^\top\bar{\mathbf{h}}}} & \text{si} \quad \bar{\mathbf{h}} \neq 0 \\ 0 & \text{si} \quad \bar{\mathbf{h}} = 0 \end{cases}, \quad \gamma_d^* = \bar{\mathbf{h}}^\top\mathbf{z}_d^* = -\|\bar{\mathbf{h}}\|_2.$$

Por tanto, se ha probado que

$$\gamma^* \geq \gamma_c^* \geq \gamma_d^* = -\|\bar{\mathbf{h}}\|_2.$$ □

Apéndice B.3. Demostración de la propiedad 4

La desigualdad $\gamma^* \geq \gamma_c^*$ se deriva de la propiedad 5. De $\bar{\lambda}\left(A(\frac{-\mathbf{h}}{\sqrt{\mathbf{h}^\top\mathbf{h}}}, \hat{w})\right) > 0$ y de la propiedad 1 se infiere que existe $\hat{\mathbf{c}} \in C(\hat{w})$ tal que

$$\hat{\mathbf{c}}^\top\left(-\frac{\mathbf{h}}{\sqrt{\mathbf{h}^\top\mathbf{h}}}\right) > 0.$$

Esto implica que $\hat{\mathbf{c}}^\top\mathbf{h} < 0$.

La propiedad (1) garantiza que el problema (6) es equivalente a

$$\begin{array}{rll} \gamma_c^* = & \min\limits_{\mathbf{z},\gamma} & \gamma \\ & s.a. & \mathbf{h}^\top\mathbf{z} \leq \gamma, \\ & & \mathbf{c}^\top\mathbf{z} \leq \gamma, \quad \forall \mathbf{c} \in \bigcup\limits_{w \in S} C(w) \\ & & \mathbf{z}^\top\mathbf{z} \leq 1. \end{array} \tag{B.1}$$

Debido a que $\hat{\mathbf{c}} \in \bigcup\limits_{w \in S} C(w)$, resulta que

$$\begin{array}{rll} \gamma_c^* \geq \tilde{\gamma}^* = & \min\limits_{\mathbf{z},\gamma} & \gamma \\ & s.a. & \mathbf{h}^\top\mathbf{z} \leq \gamma \\ & & \hat{\mathbf{c}}^\top\mathbf{z} \leq \gamma \\ & & \mathbf{z}^\top\mathbf{z} \leq 1, \end{array} \tag{B.2}$$

debido a que el problema (B.2) se obtiene del problema (B.1) quitándole restricciones. Por otra parte, de la propiedad 8 se infiere que

$$\tilde{\gamma}^* = -\min_{\mathbf{c} \in \mathrm{Co}\{\mathbf{h}, \hat{\mathbf{c}}\}} \|\mathbf{c}\|_2 = -\min_{\lambda \in [0,1]} \|(1 - \lambda)\mathbf{h} + \lambda\hat{\mathbf{c}}\|_2.$$

En la propiedad 9 se demuestra que, denotando $\hat{\gamma} = -\|\mathbf{h}\|_2$:

(i) El valor óptimo de λ se obtiene en $\lambda^* = \frac{(\mathbf{h}-\hat{\mathbf{c}})^\top\mathbf{h}}{\|\mathbf{h}-\hat{\mathbf{c}}\|_2^2} \in [0, 1]$.

(ii) $\tilde{\gamma}^* = -\|(1 - \lambda^*)\mathbf{h} + \lambda^*\hat{\mathbf{c}}\|_2 \geq \hat{\gamma}\left(1 - \frac{\hat{\gamma}^2}{8\sigma^2}\right) \geq \hat{\gamma}$.

Esto concluye la demostración. □

Apéndice B.4. Prueba del teorema 3

La demostración sigue los mismos argumentos que un resultado similar presentado en Oishi (2003). Supóngase que en la iteración k, $\mathbf{z}_k^*$ no es una solución robusta factible (de nivel δ). Es decir, $\mathrm{Prob}\,\{\, w \in W \;:\; \bar{\lambda}(A(\mathbf{z}_k^*, w)) \geq 0 \,\} > \delta$. Entonces, la probabilidad de clasificar (erróneamente) $\mathbf{z}_k^*$ como una solución robusta factible (de nivel δ) es menor que

$$(1 - \delta)^{N_k} \leq \frac{6\beta}{\pi^2(k + 1)^2}.$$

Por tanto, la probabilidad de que el algoritmo termine con una una solución robusta factible (de nivel δ) clasificada erróneamente es menor o igual que

$$\sum_{k=0}^{k_{max}} \frac{6\beta}{\pi^2(k + 1)^2} < \left(\frac{6\beta}{\pi^2}\right)\sum_{k=1}^{\infty} \frac{1}{k^2} = \left(\frac{6\beta}{\pi^2}\right)\left(\frac{\pi^2}{6}\right) = \beta.$$ □