

# SeDuMi: ejemplos introductorios

Héctor Manuel Mora Escobar

hectormora@yahoo.com

Bogotá, febrero de 2009

Esta breve introducción a SeDuMi está basada principalmente en el documento *Use SeDuMi to Solve LP, SDP and SOCP Problems: Remarks and Examples*, de Wu-Sheng Lu, de University of Victoria, Dept. of Electrical and Computer Engineering  
[menet.math.ecnu.edu.cn/lu/supplementary%20part/SeDuMi-Remarks.pdf](http://menet.math.ecnu.edu.cn/lu/supplementary%20part/SeDuMi-Remarks.pdf)

SeDuMi, self dual minimization, hecho en Matlab, implementa una técnica autotual sobre conos homogéneos autoduales. Sirve para problemas de

- optimización lineal (OL o LP),
- optimización cónica de segundo orden (SOCP)
- optimización semidefinida (SDP).

Se puede descargar en la página

<http://sedumi.ie.lehigh.edu/>

de Cor@l, Computational Optimization Research de Lehigh University en Bethlehem, Pennsylvania, E.U.A.

Una vez descargado y descomprimido el archivo, es necesario agregar al `path` de Matlab el nombre de la carpeta donde quedaron los archivos descomprimidos.

## 0.1 Optimización lineal

$$\begin{aligned} \min_x \quad & c^T x \\ & Ax = b \\ & x \geq 0, \end{aligned}$$

donde  $c, x \in \mathbb{R}^{n \times 1}$ ,  $A \in \mathbb{R}^{m \times n}$ ,  $b \in \mathbb{R}^{m \times 1}$ . El llamado desde Matlab, es simplemente

```
x = sedumi(A, b, c)
```

**Ejemplo:**

```
c = [-1 -1.4 0 0 0]';  
a = [ 1 1 1 0 0; 1 2 0 1 0; 1 0 0 0 1];  
b = [400 580 300]';  
x = sedumi(a, b, c)
```

Sedumi produce los siguientes resultados:

```
SeDuMi 1.1R3 by AdvOL, 2006 and Jos F. Sturm, 1998-2003.
Alg = 2: xz-corrector, Adaptive Step-Differentiation, theta = 0.250, beta = 0.500
eqs m = 3, order n = 6, dim = 6, blocks = 1
nnz(A) = 8 + 0, nnz(ADA) = 9, nnz(L) = 6
it :      b*y          gap      delta  rate  t/tP*  t/tD*  feas cg cg prec
 0 :              1.01E-002 0.000
 1 : -4.61E+002 2.56E-003 0.000 0.2545 0.9000 0.9000  2.39 1 1 1.2E+000
 2 : -4.65E+002 5.41E-004 0.000 0.2115 0.9000 0.9000  1.59 1 1 1.9E-001
 3 : -4.72E+002 1.17E-004 0.000 0.2164 0.9000 0.9000  1.12 1 1 4.0E-002
 4 : -4.72E+002 2.56E-006 0.000 0.0218 0.9900 0.9900  1.01 1 1
iter seconds digits      c*x          b*y
 4      0.4      Inf -4.7200000000e+002 -4.7200000000e+002
|Ax-b| = 9.5e-014, [Ay-c]_+ = 4.7E-019, |x| = 3.0e+002, |y| = 7.2e-001
```

```
Detailed timing (sec)
  Pre      IPM      Post
2.496E-001  3.588E-001  7.800E-002
Max-norms: ||b||=580, ||c|| = 1.400000e+000,
Cholesky |add|=0, |skip| = 0, ||L.L|| = 1.37397.
```

x =

```
(1,1)    220.0000
(2,1)    180.0000
(5,1)    80.0000
```

Si el llamado es `[x, y, info] = sedumi(a, b, c)`, se obtiene además

y =

```
-0.6000
-0.4000
 0.0000
```

info =

```
iter: 4
feasratio: 1
pinf: 0
dinf: 0
numerr: 0
timing: [0 0.0780 0.0312]
cpusec: 0.1092
```

## 0.2 Optimización cónica de segundo orden, SOCP

En la notación de [LVBL98] un problema de SOCP se puede escribir

$$\begin{aligned} \min \quad & f^T x \\ \text{s.t.} \quad & \|A^i x + b^i\| \leq c^{iT} x + d_i \quad i = 1, \dots, p, \end{aligned}$$

donde

$$\begin{aligned} f &\in \mathbb{R}^{n \times 1} \\ x &\in \mathbb{R}^{n \times 1} \\ A^i &\in \mathbb{R}^{(m_i-1) \times n}, \quad i = 1, \dots, p, \\ b^i &\in \mathbb{R}^{(m_i-1) \times 1}, \quad i = 1, \dots, p, \\ c^i &\in \mathbb{R}^{n \times 1}, \quad i = 1, \dots, p, \\ d &\in \mathbb{R}^p \end{aligned}$$

Con estos datos se construyen la matriz  $\tilde{A}$  y los vectores columna  $\tilde{b}$  y  $\tilde{c}$ :

$$\begin{aligned} m &= m_1 + m_2 + \dots + m_p \\ \tilde{A} &= \begin{bmatrix} -c^1 & -A^{1T} & -c^2 & -A^{2T} & \dots & -c^p & -A^{pT} \end{bmatrix} \in \mathbb{R}^{n \times m} \\ \tilde{c} &= \begin{bmatrix} d_1 \\ b^1 \\ d_2 \\ b^2 \\ \vdots \\ b_p \\ c^p \end{bmatrix} \in \mathbb{R}^{m \times 1} \\ \tilde{b} &= -f \in \mathbb{R}^{n \times 1} \end{aligned}$$

En este caso los datos para utilizar `sedumi` son:  $\tilde{A}$ ,  $\tilde{b}$ ,  $\tilde{c}$  y una estructura `K` donde se da información sobre el cono. Para este ejemplo, basta con

$$K.q = [ m_1 \quad m_2 \quad \dots \quad m_p ]$$

Si `Am`, `bm`, `cm` son matrices de Matlab con la información de  $\tilde{A}$ ,  $\tilde{b}$  y  $\tilde{c}$ , entonces el llamado es:

$$[w, x, info] = \text{sedumi}( Am, bm, cm, K)$$

**Ejemplo:**

$$\begin{aligned} \min \quad & [0 \ 0 \ 0 \ 0 \ 1] x \\ \left\| \begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} -1 \\ -3 \end{bmatrix} \right\| & \leq \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}^T x + 3 \\ \left\| \begin{bmatrix} 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 3 & 0 \end{bmatrix} x + \begin{bmatrix} -6 \\ -6 \end{bmatrix} \right\| & \leq \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}^T x + 2 \\ \left\| \begin{bmatrix} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\| & \leq \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}^T x + 0 \end{aligned}$$

Este problema de optimización cónica de segundo orden, se puede resolver mediante:

$$f = [0 \ 0 \ 0 \ 0 \ 1]';$$

$$A1 = [2 \ -1 \ 0 \ 0 \ 0; \ 0 \ 3 \ 0 \ 0 \ 0];$$

$$b1 = [-1; \ -3];$$

$$c1 = [0 \ 0 \ 0 \ 0 \ 0]';$$

$$d1 = 3;$$

$$A2 = [0 \ 0 \ 1 \ -2 \ 0; \ 0 \ 0 \ 0 \ 3 \ 0];$$

$$b2 = [-6; \ -6];$$

$$c2 = [0 \ 0 \ 0 \ 0 \ 0]';$$

$$d2 = 2;$$

$$A3 = [1 \ 0 \ -1 \ 0 \ 0; \ 0 \ 1 \ 0 \ -1 \ 0];$$

$$b3 = [0; \ 0];$$

$$c3 = [0 \ 0 \ 0 \ 0 \ 1]';$$

$$d3 = 0;$$

$$Am = [-c1 \ -A1' \ -c2 \ -A2' \ -c3 \ -A3'];$$

$$bm = -f;$$

```
cm = [ d1; b1; d2; b2; d3; b3 ];  
K.q = [ size(A1,1)+1 size(A2,1)+1 size(A3,1)+1 ];  
[w, x, info] = sedumi( Am, bm, cm, K)
```

El resultado es:

```
x =  
2.5803  
1.3470  
7.5965  
1.6232  
5.0238
```

### 0.3 Restricciones cuadráticas

Consideremos una restricción cuadrática convexa

$$x^T H x + p^T x + \beta \leq 0, \quad (1)$$

donde  $p$  y  $x$  son vectores columna y  $H \succeq 0$  (matriz simétrica semidefinida positiva) y los tamaños son compatibles.

Supongamos, adicionalmente que  $H \succ 0$  ( $H$  definida positiva). Los puntos que cumplen (1) corresponden a un elipsoide con centro en  $\bar{x}$  de la forma

$$(x - \bar{x})^T H (x - \bar{x}) \leq \rho^2 \quad (2)$$

Al desarrollar (2) se obtiene

$$x^T H x - 2(H\bar{x})^T x + \bar{x}^T H \bar{x} - \rho^2 \leq 0.$$

Entonces

$$\begin{aligned} -2H\bar{x} &= p \\ \bar{x}^T H \bar{x} - \rho^2 &= \beta \end{aligned}$$

Las igualdades anteriores permiten pasar de  $\bar{x}, \rho^2$  a  $p, \beta$  y viceversa.

Veamos ahora cómo convertir la restricción cuadrática en una restricción cónica de segundo orden. Como  $H \succ 0$ , entonces existe  $J$  tal que

$$J^T J = H.$$

Esta matriz puede ser la de la factorización de Cholesky o la raíz cuadrada de  $H$  ( $RR = H$ ) u otra matriz. Sea

$$q = \frac{1}{2}(J^{-1})^T p,$$

es decir,

$$p = 2J^T q.$$

Entonces

$$\begin{aligned}
x^T H x + p^T x + \beta &\leq 0 \\
x^T J^T J x + 2q^T J x + q^T q - q^T q + \beta &\leq 0 \\
(x^T J^T J x + 2q^T J x + q^T q) - \frac{1}{2} \left( (J^{-1})^T p \right)^T \frac{1}{2} (J^{-1})^T p + \beta &\leq 0 \\
(Jx + q)^T (Jx + q) - \frac{1}{4} p^T J^{-1} (J^{-1})^T p + \beta &\leq 0 \\
\|Jx + q\|^2 - \frac{1}{4} p^T H^{-1} p + \beta &\leq 0 \\
\|Jx + q\|^2 &\leq \frac{1}{4} p^T H^{-1} p - \beta
\end{aligned}$$

En resumen, la restricción cuadrática convexa

$$x^T H x + p^T x + \beta \leq 0,$$

es equivalente a

$$\|Jx + q\| \leq \left( \frac{1}{4} p^T H^{-1} p - \beta \right)^{1/2} \quad (3)$$

con

$$\begin{aligned}
J^T J &= H \\
q &= \frac{1}{2} (J^{-1})^T p.
\end{aligned}$$

### 0.3.1 Una restricción cuadrática particular

En el caso anterior se supuso que  $H \succ 0$ . Consideremos el caso en que  $H$  y  $p$  tienen una estructura por bloques semejante a la siguiente

$$H = \begin{bmatrix} 0 & 0 & 0 \\ 0 & \bar{H} & 0 \\ 0 & 0 & 0 \end{bmatrix}, \quad p = \begin{bmatrix} 0 \\ \bar{p} \\ 0 \end{bmatrix}$$

donde los ceros son submatrices nulas de tamaños compatibles y  $\bar{H} \succ 0$ . Obviamente  $H \succeq 0$  pero  $H \not\succeq 0$ . Sea  $H \in \mathbb{R}^{n \times n}$  y  $\bar{H} \in \mathbb{R}^{\bar{n} \times \bar{n}}$ . Si bien se puede encontrar  $\bar{J} \in \mathbb{R}^{\bar{n} \times \bar{n}}$  tal que

$$\bar{J}^T \bar{J} = \bar{H} \quad (4)$$

y así obtener  $J \in \mathbb{R}^{n \times n}$  tal que  $J^T J = H$ , no se puede obtener  $q = \frac{1}{2} (J^{-1})^T p$ .

Cálculos semejantes a los que permitieron llegar a (3) muestran que en este caso la restricción cuadrática convexa  $x^T H x + p^T x + \beta$  es equivalente a

$$\|Jx + q\| \leq \left( \frac{1}{4} \bar{p}^T \bar{H}^{-1} \bar{p} - \beta \right)^{1/2} \quad (5)$$

con

$$J = \begin{bmatrix} 0 & \bar{J} & 0 \end{bmatrix}, \quad (6)$$

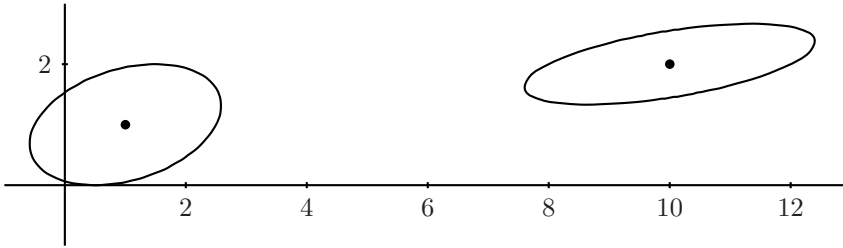
$$q = \frac{1}{2} (\bar{J}^{-1})^T \bar{p}. \quad (7)$$

Obsérvese que ahora  $J$  es una matriz rectangular en  $\mathbb{R}^{\bar{n} \times n}$  y no cuadrada en  $\mathbb{R}^{n \times n}$ . El vector  $q$  está en  $\mathbb{R}^{\bar{n} \times 1}$ .

**Ejemplo.** Hallar la distancia mínima entre las dos elipses de la forma  $(x - c)^T H(x - c) \leq \rho^2$ , donde

$$H^1 = \begin{bmatrix} 4 & -2 \\ -2 & 10 \end{bmatrix}, \quad c^1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \rho_1^2 = 9,$$

$$H^2 = \begin{bmatrix} 1 & -2 \\ -2 & 13 \end{bmatrix}, \quad c^2 = \begin{bmatrix} 10 \\ 2 \end{bmatrix}, \quad \rho_2^2 = 4.$$



Sean  $(x_1, x_2)$  las coordenadas de un punto en la primera elipse y  $(x_3, x_4)$  las coordenadas de un punto en la segunda elipse, entonces

$$\begin{aligned} & \min \left( (x_1 - x_3)^2 + (x_2 - x_4)^2 \right)^{1/2} \\ & ([x_1 \ x_2]^T - c^1)^T H^1 ([x_1 \ x_2]^T - c^1) - \rho_1^2 \leq 0 \\ & ([x_3 \ x_4]^T - c^2)^T H^2 ([x_3 \ x_4]^T - c^2) - \rho_2^2 \leq 0. \end{aligned}$$

Para quitar la función objetivo no lineal, se introduce una nueva variable  $x_5$ , cota superior para la distancia



$$\begin{aligned}
& \min x_5 \\
& ([x_1 \ x_2]^T - c^1)^T H^1 ([x_1 \ x_2]^T - c^1) - \rho_1^2 \leq 0 \\
& ([x_3 \ x_4]^T - c^2)^T H^2 ([x_3 \ x_4]^T - c^2) - \rho_2^2 \leq 0. \\
& ((x_1 - x_3)^2 + (x_2 - x_4)^2)^{1/2} \leq x_5
\end{aligned}$$

Las dos primeras restricciones son cuadráticas convexas que se pueden convertir en cónicas de segundo orden. La última es una restricción cónica de segundo orden. De manera matricial, con  $x = [x_1 \ x_2 \ x_3 \ x_4 \ x_5]^T$  se tiene

$$\begin{aligned}
& \min [0 \ 0 \ 0 \ 0 \ 1]x \\
& x^T \begin{bmatrix} 4 & -2 & 0 & 0 & 0 \\ -2 & 10 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} -4 \\ -16 \\ 0 \\ 0 \\ 0 \end{bmatrix}^T x + 1 \leq 0 \\
& x^T \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & -2 & 13 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \\ -12 \\ -12 \\ 0 \end{bmatrix}^T x + 68 \leq 0 \\
& \left\| \begin{bmatrix} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\| \leq \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}^T x + 0
\end{aligned}$$

Mediante (5), (6) y (7) las dos primeras restricciones se convierten en cónicas y el problema queda

$$\begin{aligned}
& \min [0 \ 0 \ 0 \ 0 \ 1] x \\
& \left\| \begin{bmatrix} 2 & -1 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \end{bmatrix} x + \begin{bmatrix} -1 \\ -3 \end{bmatrix} \right\| \leq \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}^T x + 3 \\
& \left\| \begin{bmatrix} 0 & 0 & 1 & -2 & 0 \\ 0 & 0 & 0 & 3 & 0 \end{bmatrix} x + \begin{bmatrix} -6 \\ -6 \end{bmatrix} \right\| \leq \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}^T x + 2 \\
& \left\| \begin{bmatrix} 1 & 0 & -1 & 0 & 0 \\ 0 & 1 & 0 & -1 & 0 \end{bmatrix} x + \begin{bmatrix} 0 \\ 0 \end{bmatrix} \right\| \leq \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix}^T x + 0
\end{aligned}$$

Este ejemplo es exactamente el ejemplo de SOCP resuelto con SeDuMi. Entonces la distancia mínima entre las dos elipses es 5.0238 obtenida entre los puntos (2.5803, 1.3470) en la primera elipse y (7.5965, 1.6232) en la segunda elipse.

## 0.4 Optimización semidefinida

Una de las formas “estándar” de problemas de SDP es la de SDPLIB, [www.nmt.edu/~sdplib](http://www.nmt.edu/~sdplib), página donde hay una colección de problemas de prueba para SDP.

$$\begin{aligned} \min \quad & c^T y \\ -A^0 + \sum_{i=1}^m y_i A^i \quad & \succcurlyeq 0 \end{aligned}$$

donde

$$\begin{aligned} c & \in \mathbb{R}^{m \times 1} \\ y & \in \mathbb{R}^{m \times 1} \\ A^i & \in \mathbb{R}^{p \times p} \text{ y es simétrica, } \quad i = 0, 1, \dots, m. \end{aligned}$$

Con estos datos se construyen matrices  $\tilde{A}$ ,  $\tilde{b}$  y  $\tilde{c}$

$$\begin{aligned} \tilde{b} & = -c \\ \tilde{c} & = -\text{vec}(A^0) \\ \tilde{A}(:, i) & = -\text{vec}(A^i), \quad i = 1, \dots, m \end{aligned}$$

Además, para el cono  $K$ ,

$$K \cdot s = p$$

Si  $AM$ ,  $bm$ ,  $cm$  son matrices de Matlab con la información de  $\tilde{A}$ ,  $\tilde{b}$  y  $\tilde{c}$ , entonces el llamado es:

$$[x, y, \text{info}] = \text{sedumi}(AM, bm, cm, K)$$

**Ejemplo.**

$$\begin{aligned} \min \quad & 10y_1 + 20y_2 \\ - \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} + y_1 \begin{bmatrix} 40 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + y_2 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 2 \\ 0 & 0 & 2 & 6 \end{bmatrix} \quad & \succcurlyeq 0 \end{aligned}$$

En Matlab:

```

c = [ 10 20]';

A0 = [1 0 0 0; 0 2 0 0; 0 0 3 0; 0 0 0 4];
A1 = [40 0 0 0; 0 10 0 0; 0 0 0 0; 0 0 0 0];
A2 = [0 0 0 0; 0 1 0 0 ; 0 0 5 2; 0 0 2 6];

bm = -c;

cm = -vec(A0)

AM(:,1) = -vec(A1);
AM(:,2) = -vec(A2);

K.s = size(A0,1)

[x, y, info] = sedumi(AM, bm, cm, K)

```

La solución es

$$y = ( 0.1, 1 ).$$

## 0.5 Optimización semidefinida con restricciones lineales

$$\begin{aligned} \min \quad & c^T y \\ & My \geq p \\ & -A^0 + \sum_{i=1}^m y_i A^i \succcurlyeq 0 \end{aligned}$$

donde

$$\begin{aligned} c &\in \mathbb{R}^{m \times 1} \\ y &\in \mathbb{R}^{m \times 1} \\ M &\in \mathbb{R}^{q \times m} \\ p &\in \mathbb{R}^{q \times 1} \\ A^i &\in \mathbb{R}^{p \times p} \text{ y es simétrica, } i = 0, 1, \dots, m. \end{aligned}$$

Con estos datos se construyen matrices  $\tilde{A}$ ,  $\tilde{b}$  y  $\tilde{c}$

$$\begin{aligned} \tilde{b} &= -c \\ \tilde{c} &= \begin{bmatrix} -p \\ -\text{vec}(A^0) \end{bmatrix} \\ \bar{A}(:, i) &= -\text{vec}(A^i), \quad i = 1, \dots, m \\ \tilde{A} &= \begin{bmatrix} -M \\ \bar{A} \end{bmatrix} \end{aligned}$$

Además, para el cono  $K$ ,

$$\begin{aligned} K.s &= p \\ K.l &= q \end{aligned}$$

Si  $AM$ ,  $bm$ ,  $cm$  son matrices de Matlab con la información de  $\tilde{A}$ ,  $\tilde{b}$  y  $\tilde{c}$ , entonces el llamado es:

$$[x, y, \text{info}] = \text{sedumi}(AM, bm, cm, K)$$

**Ejemplo.**

$$\begin{aligned}
 & \min 10y_1 + 20y_2 \\
 & \quad 4y_1 + y_2 \geq 3.6 \\
 & - \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 2 & 0 & 0 \\ 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 4 \end{bmatrix} + y_1 \begin{bmatrix} 40 & 0 & 0 & 0 \\ 0 & 10 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} + y_2 \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 5 & 2 \\ 0 & 0 & 2 & 6 \end{bmatrix} \preceq 0
 \end{aligned}$$

En Matlab:

```

c = [ 10 20]';

M = [ 4 1];
p = [ 3.6];

A0 = [1 0 0 0; 0 2 0 0; 0 0 3 0; 0 0 0 4];
A1 = [40 0 0 0; 0 10 0 0; 0 0 0 0; 0 0 0 0];
A2 = [0 0 0 0; 0 1 0 0; 0 0 5 2; 0 0 2 6];

bm = -c;

cm = [-p; -vec(A0) ]

AM(:,1) = -vec(A1);
AM(:,2) = -vec(A2);
AM = [-M; AM]

K.s = size(A0,1);
K.l = size(M,1);

[x, y, info] = sedumi(AM, bm, cm, K)

```

La solución es

$$y = ( 0.65, 1 ).$$

## 0.6 Restricciones cónicas de segundo orden y SDP

Se puede mostrar que

$$\|u\| \leq t \Rightarrow \begin{bmatrix} tI & u \\ u^T & t \end{bmatrix} \succcurlyeq 0$$

**Ejemplo:** (tomado de Cornuejols, p. 183) Se conoce una matriz  $\hat{\Sigma}$  simétrica, con unos en la diagonal y se supone que es una aproximación de una matriz de correlaciones pero no es semidefinida positiva. Se desea encontrar una matriz  $\Sigma$  parecida a  $\hat{\Sigma}$  (en el sentido de mínimos cuadrados) que sea semidefinida positiva.

Sean

$$\hat{\Sigma} = \begin{bmatrix} 1.0 & 0.8 & 0.5 & 0.2 \\ 0.8 & 1.0 & 0.9 & 0.1 \\ 0.5 & 0.9 & 1.0 & 0.7 \\ 0.2 & 0.1 & 0.7 & 1.0 \end{bmatrix}, \quad \Sigma = \begin{bmatrix} 1.0 & y_1 & y_2 & y_3 \\ y_1 & 1.0 & y_4 & y_5 \\ y_2 & y_4 & 1.0 & y_6 \\ y_3 & y_5 & y_6 & 1.0 \end{bmatrix}.$$

Se desea resolver el siguiente problema de optimización:

$$\begin{aligned} \min \quad & \|\Sigma - \hat{\Sigma}\|_F \\ & \Sigma \succcurlyeq 0. \end{aligned}$$

La norma  $\|\cdot\|_F$  es la norma de Frobenius,

$$\|M\|_F = \left( \sum_i \sum_j m_{ij}^2 \right)^{1/2}.$$

La función objetivo, que no es lineal, se puede convertir en lineal mediante la introducción de una nueva variable, cota superior para  $\|\Sigma - \hat{\Sigma}\|_F$ .

$$\begin{aligned} \min \quad & y_7 \\ \|\Sigma - \hat{\Sigma}\|_F & \leq y_7 \\ & \Sigma \succcurlyeq 0, \end{aligned}$$

$$\begin{aligned}
& \min y_7 \\
& (2(y_1 - 0.8)^2 + 2(y_2 - 0.5)^2 + 2(y_3 - 0.2)^2 + \\
& 2(y_4 - 0.9)^2 + 2(y_5 - 0.1)^2 + 2(y_6 - 0.7)^2)^{1/2} \leq y_7 \\
& -A_0 + \sum_{i=1}^7 y_i A_i \succcurlyeq 0,
\end{aligned}$$

donde

$$\begin{aligned}
A_0 = -I, \quad A_1 = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad A_2 = \begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad A_3 = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 \end{bmatrix}, \\
A_4 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}, \quad A_5 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}, \quad A_6 = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{bmatrix}, \quad A_7 = 0.
\end{aligned}$$

Entonces

$$\begin{aligned}
& \min y_7 \\
& \|\tilde{y} - \sigma\| \leq ky_7 \\
& -A_0 + \sum_{i=1}^7 y_i A_i \succcurlyeq 0,
\end{aligned}$$

con

$$k = \frac{1}{\sqrt{2}}, \quad \tilde{y} = [y_1 \quad y_2 \quad y_3 \quad y_4 \quad y_5 \quad y_6]^\top, \quad \sigma = [0.8 \quad 0.5 \quad 0.2 \quad 0.9 \quad 0.1 \quad 0.7]^\top,$$

$$\begin{aligned}
& \min y_7 \\
& \begin{bmatrix} ky_7 I_6 & \tilde{y} - \sigma \\ \tilde{y}^\top - \sigma^\top & ky_7 \end{bmatrix} \succcurlyeq 0 \\
& -A_0 + \sum_{i=1}^7 y_i A_i \succcurlyeq 0,
\end{aligned}$$



$$\begin{aligned} \min \quad & y_7 \\ -B_0 + \sum_{i=1}^7 y_i B_i \geq & 0, \\ -A_0 + \sum_{i=1}^7 y_i A_i \geq & 0, \end{aligned}$$

donde

$$B_0 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0.8 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.5 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.9 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0.7 \\ 0.8 & 0.5 & 0.2 & 0.9 & 0.1 & 0.7 & 0 \end{bmatrix}, \quad B_1 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix},$$

$$\dots, B_6 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 \end{bmatrix}, \quad B_7 = k I_7.$$

$$\begin{aligned} \min \quad & y_7 \\ -A_0 + \sum_{i=1}^7 y_i A_i \geq & 0, \end{aligned}$$

con

$$A_i = \begin{bmatrix} B_i & 0 \\ 0 & A_i \end{bmatrix}, \quad i = 0, \dots, 7.$$

En Matlab

```
c = [ 0 0 0 0 0 0 0 1]';

A0 = -eye(4,4);
A1 = zeros(4,4); A1(1,2) = 1; A1(2,1) = 1;
A2 = zeros(4,4); A2(1,3) = 1; A2(3,1) = 1;
```

```

A3 = zeros(4,4); A3(1,4) = 1; A3(4,1) = 1;
A4 = zeros(4,4); A4(2,3) = 1; A4(3,2) = 1;
A5 = zeros(4,4); A5(2,4) = 1; A5(4,2) = 1;
A6 = zeros(4,4); A6(3,4) = 1; A6(4,3) = 1;
A7 = zeros(4,4);

k = 1/sqrt(2);
sigma = [ 0.8 0.5 0.2 0.9 0.1 0.7];

B0 = zeros(7,7); B0(1:6,7) = sigma; B0(7,1:6) = sigma;
B1 = zeros(7,7); B1(1,7) = 1; B1(7,1) = 1;
B2 = zeros(7,7); B2(2,7) = 1; B2(7,2) = 1;
B3 = zeros(7,7); B3(3,7) = 1; B3(7,3) = 1;
B4 = zeros(7,7); B4(4,7) = 1; B4(7,4) = 1;
B5 = zeros(7,7); B5(5,7) = 1; B5(7,5) = 1;
B6 = zeros(7,7); B6(6,7) = 1; B6(7,6) = 1;
B7 = k*eye(7,7);

AA0 = [B0 zeros(7,4);zeros(4,7) A0];
AA1 = [B1 zeros(7,4);zeros(4,7) A1];
AA2 = [B2 zeros(7,4);zeros(4,7) A2];
AA3 = [B3 zeros(7,4);zeros(4,7) A3];
AA4 = [B4 zeros(7,4);zeros(4,7) A4];
AA5 = [B5 zeros(7,4);zeros(4,7) A5];
AA6 = [B6 zeros(7,4);zeros(4,7) A6];
AA7 = [B7 zeros(7,4);zeros(4,7) A7];

bm = -c;

cm = -vec(AA0);

AM(:,1) = -vec(AA1);
AM(:,2) = -vec(AA2);
AM(:,3) = -vec(AA3);
AM(:,4) = -vec(AA4);
AM(:,5) = -vec(AA5);
AM(:,6) = -vec(AA6);
AM(:,7) = -vec(AA7);

K.s = size(AA0,1);

[x, y, info] = sedumi(AM, bm, cm, K)

```

La solución es:

$$y = (0.7710, 0.5291, 0.1835, 0.8153, 0.1480, 0.6517, 0.1660)$$

Entonces

$$\Sigma = \begin{bmatrix} 1 & 0.7710 & 0.5291 & 0.1835 \\ 0.7710 & 1 & 0.8153 & 0.1480 \\ 0.5291 & 0.8153 & 1 & 0.6517 \\ 0.1835 & 0.1480 & 0.6517 & 1 \end{bmatrix}$$

$$\|\Sigma - \hat{\Sigma}\|_F = y_7 = 0.1660.$$

**Ejemplo:** La matriz de correlaciones resultado del ejemplo anterior es singular o casi singular,  $\det(\Sigma) = -1.6700 \times 10^{-9}$  (el valor propio mínimo es nulo o casi nulo), lo que indica redundancia entre las variables explicativas. Se desea ahora que el valor propio mínimo de  $\Sigma$  sea mayor o igual a 0.2. Esta restricción adicional se puede escribir

$$\Sigma - tI \succcurlyeq 0.$$

con  $t = 0.2$ . El problema resultante es

$$\begin{aligned} \min \quad & y_7 \\ \|\Sigma - \hat{\Sigma}\|_F \leq & y_7 \\ \Sigma - tI \succcurlyeq & 0. \end{aligned}$$

En el ejemplo anterior  $A_0 = -I$ , ahora  $A_0 = -(1-t)I$ .

$$c = [0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 0 \ 1]';$$

$$t = 0.2;$$

$$A0 = -(1-t)*eye(4,4);$$

$$A1 = zeros(4,4); A1(1,2) = 1; A1(2,1) = 1;$$

$$A2 = zeros(4,4); A2(1,3) = 1; A2(3,1) = 1;$$

$$A3 = zeros(4,4); A3(1,4) = 1; A3(4,1) = 1;$$

$$A4 = zeros(4,4); A4(2,3) = 1; A4(3,2) = 1;$$

$$A5 = zeros(4,4); A5(2,4) = 1; A5(4,2) = 1;$$

$$A6 = zeros(4,4); A6(3,4) = 1; A6(4,3) = 1;$$

$$A7 = zeros(4,4);$$

$$k = 1/\sqrt{2};$$

$$\text{sigma} = [0.8 \ 0.5 \ 0.2 \ 0.9 \ 0.1 \ 0.7];$$

$$B0 = zeros(7,7); B0(1:6,7) = \text{sigma}; B0(7,1:6) = \text{sigma};$$

$$B1 = zeros(7,7); B1(1,7) = 1; B1(7,1) = 1;$$

$$B2 = zeros(7,7); B2(2,7) = 1; B2(7,2) = 1;$$

$$B3 = zeros(7,7); B3(3,7) = 1; B3(7,3) = 1;$$

```

B4 = zeros(7,7); B4(4,7) = 1; B4(7,4) = 1;
B5 = zeros(7,7); B5(5,7) = 1; B5(7,5) = 1;
B6 = zeros(7,7); B6(6,7) = 1; B6(7,6) = 1;
B7 = k*eye(7,7);

AA0 = [B0 zeros(7,4);zeros(4,7) A0];
AA1 = [B1 zeros(7,4);zeros(4,7) A1];
AA2 = [B2 zeros(7,4);zeros(4,7) A2];
AA3 = [B3 zeros(7,4);zeros(4,7) A3];
AA4 = [B4 zeros(7,4);zeros(4,7) A4];
AA5 = [B5 zeros(7,4);zeros(4,7) A5];
AA6 = [B6 zeros(7,4);zeros(4,7) A6];
AA7 = [B7 zeros(7,4);zeros(4,7) A7];

bm = -c;

cm = -vec(AA0);

AM(:,1) = -vec(AA1);
AM(:,2) = -vec(AA2);
AM(:,3) = -vec(AA3);
AM(:,4) = -vec(AA4);
AM(:,5) = -vec(AA5);
AM(:,6) = -vec(AA6);
AM(:,7) = -vec(AA7);

K.s = size(AA0,1);

[x, y, info] = sedumi(AM, bm, cm, K)

Sigma = [1 y(1) y(2) y(3); y(1) 1 y(4) y(5); y(2) y(4) 1 y(6); y(3) y(5) y(6) 1]
v_prop = eig(Sigma)

```

La solución es:

$$y = (0.7341, 0.5673, 0.1648, 0.6783, 0.2161, 0.5815, 0.4166)$$

Entonces

$$\Sigma = \begin{bmatrix} 1 & 0.7341 & 0.5673 & 0.1648 \\ 0.7341 & 1 & 0.6783 & 0.2161 \\ 0.5673 & 0.6783 & 1 & 0.5815 \\ 0.1648 & 0.2161 & 0.5815 & 1 \end{bmatrix}$$

$$\|\Sigma - \hat{\Sigma}\|_F = y_7 = 0.4166,$$

valores propios : 0.2000, 0.2949, 0.9794, 2.5257.

## Referencias

[CoT07] Cornuejols y Tütüncü, *Optimizaton Methods in Finance*, Cambridge University Press, 2007.

[LVBL98] Lobo, Vandenberghe, Boyd y Lebret, *Applications of second-Order cone programming*, Linear Algebra and its Applications 284 (1998).