Julie Eaton
Convex Optimization

Chapter 4, problem 43
Eigenvalue optimization via SDP. Suppose $A : \mathbb{R}^n \to \mathbb{S}^m$ is affine, *i.e.*,
$A(x) = A_0 + x_1A_1 + \cdots + x_nA_n$ where $A_i \in \mathbb{S}^m$. Let $\lambda_1(x) \geq \cdots \geq \lambda_m(x)$ denote the eigenvalues of $A(x)$. Show how to pose the following problems as SDPs.

Target form for the SDP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & x_1F_1 + \ldots x_nF_n + G \preceq 0 \\ & Ax = b \end{array}$$

where $F_1, \ldots, F_n, G \in \mathbb{S}^k$ and $A \in \mathbb{R}^{p\times n}$.
Generalized Lagrangian Duality

Primal Problem	$(\mathcal{P})$	$\min_{x \text{ s.t. } f_i(x)\preceq_K 0, i=1,\ldots m} f_0(x)$
Lagrangian		$L(x,y) = f_0(x) + \langle y, F(x)\rangle - \delta_K(y)$
Dual objective		$g(y) = \inf_x L(x,y).$
Dual Problem	$(\mathcal{D})$	$\sup_y g(y)$

(a) Minimize the maximum eigenvalue $\lambda_1(x)$.
Solution: Note that $\lambda I - A(x) \succeq 0$ if and only if $\lambda - \lambda_1(x) \geq 0$. Therefore, $\lambda_1(x)$ is the smallest λ for which $\lambda I - A(x) \succeq 0$. That is, for each x we have

$$\lambda_1(x) = \min_{A(x)-\lambda I\preceq 0} \lambda \tag{1}$$

So we may express the problem at hand as follows $\min_{x\in\mathbb{R}^n} \lambda_1(x) = \min_{x\in\mathbb{R}^n \text{ s.t. } A(x)-\lambda I\preceq 0} \lambda$. Let $\tilde{c} = [0,\ldots,0,1]^T \in \mathbb{R}^{n+1}$, and let $\tilde{x} = [x,\lambda]^T$. Then we may re-express the minimization problem above as

$$\min_{A_0+x_1A_1+\cdots+x_nA_n+\lambda(-I)\preceq 0} \tilde{c}^T\tilde{x},$$

which is in the primal form for an SDP.
The Lagrangian is

$$\begin{aligned} L(\tilde{x},Y) &= \tilde{c}^T\tilde{x} + \operatorname{tr}\left((A_0 + x_1A_1 + \cdots x_nA_n + \lambda(-I))Y\right) - \delta_{\mathbb{S}^m_+}(Y) \\ &= \operatorname{tr}(A_0Y) + \sum_{i=1}^n x_i\operatorname{tr}(A_iY) + \lambda(1 - \operatorname{tr}(Y)) - \delta_{\mathbb{S}^m_+}(Y) \end{aligned}$$

Therefore,

$$g(Y) = \inf_{\tilde{x}} L(\tilde{x},Y) = \begin{cases} \operatorname{tr}(A_0Y) - \delta_{\mathbb{S}^m_+}(Y) & \text{if } \operatorname{tr}(A_iY) = 0 \text{ for } i = 1,\ldots,n \text{ and } \operatorname{tr}(Y) = 1, \\ -\infty & \text{otherwise.} \end{cases}$$

So the dual problem is

$$\begin{array}{c} \max\ \operatorname{tr}(A_0Y) \\ Y \in \mathbb{S}^m_+ \text{ subject to} \\ \operatorname{tr}(Y) = 1, \operatorname{tr}(A_iY) = 0 \\ \text{for } i = 1,\ldots n \end{array}$$

(b) Minimize the spread of the eigenvalues, $\lambda_1(x) - \lambda_m(x)$.
Solution: Note that $\gamma I - A(x) \preceq 0$ if and only if $\gamma - \lambda_m(x) \leq 0$. Therefore,

$$\lambda_m(x) = \max_{\gamma I - A(x) \preceq 0} \gamma. \tag{2}$$

Negating both sides of Equation 2, and combining the expression for $\lambda_m(x)$ with that for $\lambda_1(x)$ in Equation 1, we have

$$\min_{x\in\mathbb{R}^n} \lambda_1(x) - \lambda_m(x) = \min_{x\in\mathbb{R}^n \text{ s.t. } \gamma I \preceq A(x) \preceq \lambda I} \lambda - \gamma.$$

Let $F_i = \begin{bmatrix} A_i & 0 \\ 0 & -A_i \end{bmatrix}$ for $i = 0, \dots, n$, let $F_{\lambda_1} = \begin{bmatrix} -I & 0 \\ 0 & 0 \end{bmatrix}$, and $F_{\lambda_m} = \begin{bmatrix} 0 & 0 \\ 0 & I \end{bmatrix}$.
Define $\tilde{x} = [x, \lambda, \gamma]^T$ and $\tilde{c} = [0, \dots, 0, 1, -1]^T \in \mathbb{R}^{n+2}$. Then

$$\min_{x\in\mathbb{R}^n} \lambda_1(x) - \lambda_m(x) = \min_{F_0 + \sum_{i=1}^m x_1 F_1 + \lambda F_\lambda + \gamma F_\gamma \preceq 0} \tilde{c}^T \tilde{x},$$

which is an SDP. The Lagrangian is

$$\begin{aligned} L(\tilde{x}, Y) &= \tilde{c}^T \tilde{x} + \operatorname{tr}\left((F_0 + x_1 F_1 + \cdots + x_n F_n + \lambda F_\lambda + \gamma F_\gamma) Y\right) - \delta_{\mathbb{S}^{2m}_+})(Y) \\ &= \operatorname{tr}(F_0 Y) + \sum_{i=1}^n x_i \operatorname{tr}(F_i Y) + \lambda(1 + \operatorname{tr}(F_\lambda Y)) + \gamma(-1 + \operatorname{tr}(F_\gamma Y)) - \delta_{\mathbb{S}^{2m}_+}(Y) \end{aligned}$$

Therefore,

$$g(Y) = \inf_{\tilde{x}} L(\tilde{x}, Y) = \begin{cases} \operatorname{tr}(F_0 Y) - \delta_{\mathbb{S}^{2m}_+}(Y) & \text{if } \operatorname{tr}(F_i Y) = 0 \text{ for } i = 1, \dots, n, \\ & \operatorname{tr}(F_\lambda Y) = -1, \text{ and } \operatorname{tr}(F_\gamma Y) = 1, \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem is

$$\begin{aligned} & \max\ \operatorname{tr}(F_0 Y). \\ Y \in \mathbb{S}^{2m}_+ & \text{ subject to} \\ \operatorname{tr}(F_\lambda Y) = -1, \operatorname{tr}(F_\gamma Y) = 1, \operatorname{tr}(F_i Y) = 0 & \\ \text{for } i = 1, \dots n & \end{aligned}$$

If we write $Y = \begin{bmatrix} Y_1 & \tilde{Y} \\ \tilde{Y}^T & Y_2 \end{bmatrix}$, where $Y_1, Y_2 \in \mathbb{S}^m_+$, then we may re-express the the dual problem as

$$\begin{aligned} & \max\ \operatorname{tr}(A_0 Y_1) - \operatorname{tr}(A_0 Y_2). \\ Y_1, Y_2 \in \mathbb{S}^m_+ & \text{ subject to} \\ \operatorname{tr}(Y_1) = 1, \operatorname{tr}(Y_2) = 1, \operatorname{tr}(A_i Y_1) - \operatorname{tr}(A_i Y_2) = 0 & \\ \text{for } i = 1, \dots n & \end{aligned}$$

(c) Minimize the condition number of $A(x)$, subject to $A(x) \succ 0$. The condition number is defined as $\kappa(A(x)) = \lambda_1(x)/\lambda_m(x)$. You may assume that $A(x) \succ 0$ for at least one x.
Solution: From Equation 1 and by taking the reciprocal of each side of Equation 2, we have

$$\min_{\{x\in\mathbb{R}^n : A(x) \succ 0\}} \frac{\lambda_1(x)}{\lambda_m(x)} = \min_{0 \prec \gamma I \preceq A(x) \preceq \lambda I} \frac{\lambda}{\gamma}.$$

Let $s = 1/\gamma$ and $t = \lambda/\gamma$. Then we may express the minimization problem as

$$\min_{-s<0, I \preceq sA(x) \preceq tI} t = \min_{-s<0, I \preceq sA_0 + \sum_{i=1}^n x_i s A_i \preceq tI} t.$$

Letting $y = sx$, we have

$$\min_{-s<0, I \preceq sA_0 + \sum_{i=1}^n x_i s A_i \preceq tI} t = \min_{-Is \prec 0, I \preceq sA_0 + \sum_{i=1}^n y_i A_i \preceq tI} t.$$

Consider the two optimization problems:

$$(\mathcal{P}_1) \qquad \min_{-Is \prec 0, I \preceq sA_0 + \sum_{i=1}^n y_i A_i \preceq tI} t$$

$$(\mathcal{P}_2) \qquad \min_{-Is \preceq 0, I \preceq sA_0 + \sum_{i=1}^n y_i A_i \preceq tI} t.$$

The second problem is a convex optimization problem, as defined in chapter 4. Both objectives are differentiable. The one difficulty $(\mathcal{P}_1)$ is that its domain is not closed. Note that the proof of the optimality condition (4.21) does not rely on the fact that the domain of the problem is closed–the same proof works for a problem with a domain that convex but not closed. Also, the discussion in section 4.6 says that condition (4.21) may be applied to problems with generalized inequality constraints. Therefore, we may apply this optimality condition to both $\mathcal{P}_1$ and $\mathcal{P}_2$. Suppose $[\hat{y}, \hat{s}, \hat{t}]^T$ solves $(\mathcal{P}_1)$. Then

$$[0,1,0]([y - \hat{y}, t - \hat{t}, s - \hat{s}]^T) = t - \hat{t} \geq 0 \tag{3}$$

for all $[y, t, s]^T \in \mathbb{R}^{n+2}$ such that $s > 0$ and $I \preceq sA_0 + \sum_{i=1}^n y_i A_i \preceq tI$. Then Equation 3 holds for all $[y, t, s]^T \in \mathbb{R}^{n+2}$ such that $s \geq 0$ and $I \preceq sA_0 + \sum_{i=1}^n y_i A_i \preceq tI$, which implies that $[\hat{y}, \hat{t}, \hat{s}]^T$ solves $(\mathcal{P}_2)$. So, every solution to $(\mathcal{P}_1)$ is a solution to $(\mathcal{P}_2)$. The converse is not true. In order to use the SDP framework, I will pose $(\mathcal{P}_2)$ as an SDP.

Let $G = \begin{bmatrix} 0 & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & 0 \end{bmatrix}$, $F_s = \begin{bmatrix} -I & 0 & 0 \\ 0 & -A_0 & 0 \\ 0 & 0 & A_0 \end{bmatrix}$, $F_i = \begin{bmatrix} 0 & 0 & 0 \\ 0 & -A_i & 0 \\ 0 & 0 & A_i \end{bmatrix}$, and $F_t = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -I \end{bmatrix}$.

Let $\tilde{x} = [y, t, s]^T \in \mathbb{R}^{n+2}$, and $\tilde{c} = [0, \cdots, 0, 1, 0]^T \in \mathbb{R}^{n+2}$. Then $(\mathcal{P}_2)$ is equivalent to

$$\min_{sF_s + \sum_{i=1}^n y_i F_i + tF_t + G \preceq 0} \tilde{c}^T \tilde{x}.$$

The Lagrangian is

$$\begin{aligned} L(\tilde{x}, Y) &= \tilde{c}^T \tilde{x} + \mathrm{tr}((sF_s + \sum_{i=1}^n y_i F_i + tF_t + G)Y) - \delta_{\mathbb{S}_+^{3m}}(Y) \\ &= s\mathrm{tr}(F_s Y) + \sum_{i=1^n} y_i \mathrm{tr}(F_i Y) + t(1 + \mathrm{tr}(F_t Y)) + \mathrm{tr}(GY) - \delta_{\mathbb{S}_+^{3m}}(Y) \end{aligned}$$

Therefore,

$$g(Y) = \inf_{\tilde{x}} L(\tilde{x}, Y) = \begin{cases} \mathrm{tr}(GY) - \delta_{\mathbb{S}_+^{3m}}(Y) & \text{if } \mathrm{tr}(F_i Y) = 0 \text{ for } i = 1, \ldots, n \text{ and } \mathrm{tr}(F_t Y) = -1, \\ -\infty & \text{otherwise.} \end{cases}$$

So the dual problem is

$$\max_{Y \in \mathbb{S}_+^{3m}} \mathrm{tr}(GY).$$

$$\text{subject to}$$

$$\mathrm{tr}(F_s Y) = 0, \mathrm{tr}(F_i Y) = 0, \mathrm{tr}(F_t Y) = -1$$

$$\text{for } i = 1, \ldots n$$

If we write $Y = \begin{bmatrix} Y_1 & A & B \\ A^T & Y_2 & C \\ B^T & C^T & Y_3 \end{bmatrix}$, where $Y_1, Y_2, Y_3 \in \mathbb{S}_+^m$, then we may re-express the the dual problem as

$$\max \ \mathrm{tr}(Y_2).$$

$$Y_1, Y_2, Y_3 \in \mathbb{S}_+^m \text{ subject to}$$

$$\mathrm{tr}(-Y_1 - A_0 Y_2 + A_0 Y_3) = 0, \mathrm{tr}(A_i(Y_3 - Y_2)) = 0, \mathrm{tr}(Y_3) = 1$$

$$\text{for } i = 1, \ldots n$$

(d) Minimize the sum of the absolute values of the eigenvalues, $|\lambda_1(x)| + \cdots + |\lambda_m(x)|$.
Solution: For each x, $A(x)$ is a linear mapping from $\mathbb{R}^m \to \mathbb{R}^m$. Since $A(x)$ is symmetric, it has an orthogonal basis of eigenvectors (which depend smoothly on x). Let E_+ denote the Minkowski sum of the eigenspaces of $A(x)$ corresponding to the nonnegative eigenvalues of $A(x)$ (let E_+ be 0 if there are no nonnegative eigenvalues). Likewise, let E_- denote the Minkowski sum of the eigenspaces of $A(x)$ corresponding to negative eigenvalues of $A(x)$ (and let it be 0 if all the eigenvalues are nonnegative). Let $\tilde{U} = -A(x)|_{E_+}$ and let $\tilde{V} = A(x)|_{E_-}$. Then $\tilde{U}, \tilde{V} \in \mathbb{S}_+^m$ and $A(x) = \tilde{U} - \tilde{V}$. Furthermore, the eigenvalues of $\tilde{U}$ are the nonnegative eigenvalues of $A(x)$ and the eigenvalues of $\tilde{V}$ are the absolute values of the negative eigenvalues of $A(x)$. Then $\sum_{i=1}^m |\lambda_i(x)| = \mathrm{tr}(\tilde{U} + \tilde{V})$. Suppose $U, V \in \mathbb{S}_+^m$ are such that $A(x) = U - V$. We may write $U = U_+ + U_-$ and $V = V_+ + V_-$, where $U_+ = U|_{E_+}$, etc. Note that $U_+ \succeq 0$, etc. Then $A(x) = (U_+ - V_+) - (V_- - U_-)$, which implies that $\tilde{U} = U_+ + V_+$ and $\tilde{V} = V_- - U_-$. So,

$$\begin{aligned}\sum_{i=1}^m |\lambda_i(x)| &= \mathrm{tr}(\tilde{U} + \tilde{V}) = \mathrm{tr}(U_+ - V_+ + V_- - U_-) \\ &= \mathrm{tr}(U_+) - \mathrm{tr}(V_+) + \mathrm{tr}(V_-) - \mathrm{tr}(U_-) \le \mathrm{tr}(U_+) + \mathrm{tr}(V_-) \\ &\le \mathrm{tr}(U) + \mathrm{tr}(V).\end{aligned}$$

Therefore, the problem $\min_x \sum_{i=1}^m |\lambda_i(x)|$ is equivalent to

$$\min_{U,V \succeq 0, A_0 + \sum_{i=1}^n x_i A_i = U - V} \mathrm{tr}(U + V).$$