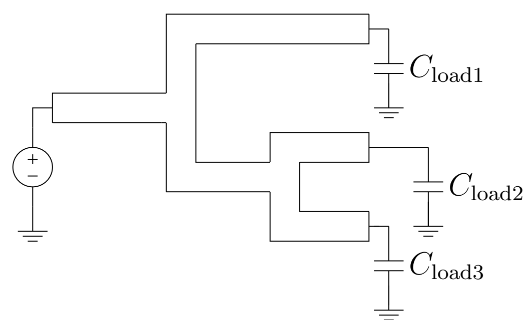
Homework 5

Submit answers for problems 1–5. Problems 6–10 are practice problems.

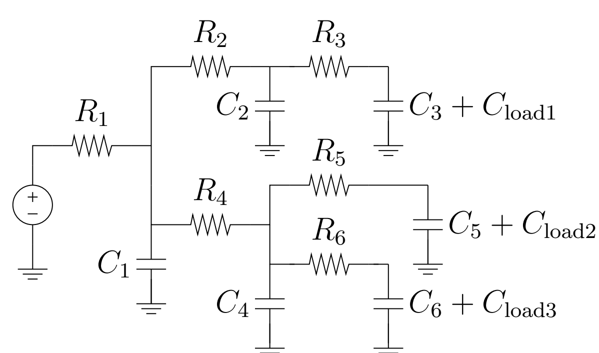
1. Exercise A16.2.

2. Exercise A8.9.

3. *Interconnect sizing.* We consider the problem of sizing the interconnecting wires of the simple circuit shown below, in which one voltage source drives three different capacitive loads $C_{\rm load1}$, $C_{\rm load2}$, and $C_{\rm load3}$.



We divide the wires into 6 segments of fixed length l_i; the optimization variables in the problem will be the widths w_i of the segments. (The height of the wires is related to the particular integrated circuit technology process, and is fixed.) We take the lengths l_i to be one, for simplicity. In the next figure each of the wire segments is modeled by a simple RC circuit, with the resistance inversely proportional to the width of the segment and the capacitance proportional to the width.



The capacitance and resistance of the ith segment is thus

$$C_i = k_0 w_i, \qquad R_i = \rho / w_i, \qquad i = 1, \ldots, 6,$$

where k_0 and ρ are positive constants, which we take to be one for simplicity. We also take $C_{\mathrm{load1}} = 1.5$, $C_{\mathrm{load2}} = 1$, and $C_{\mathrm{load3}} = 5$. We are interested in the trade-off between two objectives: area and delay. The total area used by the wires is

$$A(w) = \sum_{i=1}^{6} w_i l_i = \sum_{i=1}^{6} w_i.$$

We use the *Elmore delay* to model the delay from the source to each of the loads. The Elmore delays to loads 1, 2, and 3 are defined as

$$\begin{aligned}
T_1 &= (C_3 + C_{\mathrm{load1}})(R_1 + R_2 + R_3) + C_2(R_1 + R_2) \\
&\quad + (C_1 + C_4 + C_5 + C_6 + C_{\mathrm{load2}} + C_{\mathrm{load3}})R_1 \\
T_2 &= (C_5 + C_{\mathrm{load2}})(R_1 + R_4 + R_5) + C_4(R_1 + R_4) \\
&\quad + (C_6 + C_{\mathrm{load3}})(R_1 + R_4) + (C_1 + C_2 + C_3 + C_{\mathrm{load1}})R_1 \\
T_3 &= (C_6 + C_{\mathrm{load3}})(R_1 + R_4 + R_6) + C_4(R_1 + R_4) \\
&\quad + (C_1 + C_2 + C_3 + C_{\mathrm{load1}})R_1 + (C_5 + C_{\mathrm{load2}})(R_1 + R_4).
\end{aligned}$$

(The general rule is as follows: the Elmore delay from the source to node j is given by

$$\sum_{\text{all nodes } i} C_i R_{ij}$$

where C_i is the capacitance at node i and R_{ij} is the sum of the resistances on the intersection of the path from the source to node i and the path from the source to node j.) The second objective in the trade-off is the maximum of these delays,

$$T(w) = \max\{T_1, T_2, T_3\}.$$

We also impose minimum and maximum allowable values for the wire widths:

$$W_{\min} \le w_i \le W_{\max}, \quad i = 1, \ldots, 6.$$

For our specific problem, we take $W_{\min} = 0.1$ and $W_{\max} = 10$.

We compare two choices of wire widths.

(a) *Equal wire widths.* Plot the values of area $A(w)$ versus delay $T(w)$, obtained if you take equal wire widths w_i (varying between $W_{\min}$ and $W_{\max}$).

(b) *Optimal wire widths.* The optimal area-delay trade-off curve can be computed by scalarization, *i.e.*, by minimizing $A(w) + \mu T(w)$, subject to the constraints on w, for a large number of different positive values of μ. Verify that the scalarized problem is a geometric program (GP). For the specific problem parameters given, compute the area-delay trade-off curve using CVX or CVXPY. You can choose the values of μ logarithmically spaced between 10^{-3} and 10^3. Compare the optimal trade-off curve with the one obtained in part (a).

Consult chapter 7 of the CVX user guide for details on how to solve GPs. For reasons explained in the user guide, CVX is not very fast when solving GPs. If needed, you can limit the number of weights μ, for example, to 10 or 20.

4. Exercise A4.11 (b,c).

5. *Minimizing symmetric convex functions of eigenvalues.* If $g : \mathbf{R} \to \mathbf{R}$ is a convex function of a scalar variable, then the function $f : \mathbf{S}^m \to \mathbf{R}$ defined by

$$f(X) = \sum_{i=1}^{m} g(\lambda_i(X)), \tag{1}$$

where $\lambda_1(X), \ldots, \lambda_m(X)$ are the eigenvalues of X, is convex. This is an application of the result in exercise A3.21. In this problem we consider two specific choices for g and derive semidefinite programming formulations of the problem

$$\text{minimize} \quad f(A(y))$$

with variable $y \in \mathbf{R}^n$, where $A(y) = A_0 + y_1 A_1 + \cdots + y_n A_n$ and $A_0, \ldots, A_n \in \mathbf{S}^m$. The two examples are easily extended to other SDP-representable functions g.

(a) Define $g(x) = x^{3/2}$ with $\mathbf{dom}\, g = \mathbf{R}_+$. Verify that $g(x)$ is the optimal value of the SDP

$$\begin{array}{ll} \text{minimize} & v \\ \text{subject to} & \begin{bmatrix} x & u \\ u & 1 \end{bmatrix} \succeq 0, \quad \begin{bmatrix} u & x \\ x & v \end{bmatrix} \succeq 0, \end{array}$$

with variables $u, v \in \mathbf{R}$. Then show that $f(X) = \sum_{i=1}^{m} \lambda_i(X)^{3/2}$, with $X \in \mathbf{S}^m_+$, is the optimal value of the SDP

$$\begin{array}{ll} \text{minimize} & \mathbf{tr}\, V \\ \text{subject to} & \begin{bmatrix} X & U \\ U & I \end{bmatrix} \succeq 0, \quad \begin{bmatrix} U & X \\ X & V \end{bmatrix} \succeq 0, \end{array} \tag{2}$$

with variables $U, V \in \mathbf{S}^m$.

(b) Next, let g be the Huber penalty function

$$g(x) = \begin{cases} x^2 & |x| \leq 1 \\ 2|x| - 1 & |x| > 1. \end{cases}$$

Show that $g(x)$ is the optimal value of the SDP

$$\begin{array}{ll} \text{minimize} & v \\ \text{subject to} & \begin{bmatrix} v - u & x \\ x & u + 1 \end{bmatrix} \succeq 0, \quad u \geq 0, \end{array}$$

with variables u, v. Then show that $f(X) = \sum_{i=1}^m g(\lambda_i(X))$ is the optimal value of the SDP

$$\begin{array}{ll} \text{minimize} & \mathbf{tr}\, V \\ \text{subject to} & \left[\begin{array}{cc} V-U & X \\ X & U+I \end{array}\right], \quad U \succeq 0, \end{array} \tag{3}$$

with variables $U, V \in \mathbf{S}^m$.

The constraints in the SDPs (2) and (3) are also linear in X. To minimize $f(A(y))$ we can therefore make the substitution $X = A(y)$ and optimize jointly over y, U, V.

6. Exercise A4.11 (a,d).

Solution.

(a) We introduce a scalar variable t and write the problem as

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & f(x) \leq t. \end{array}$$

We then use the Schur complement theorem to express this as an SDP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \left[\begin{array}{cc} F(x) & c \\ c^T & t \end{array}\right] \succeq 0 \end{array}$$

with variables x, t.

Note that there is a small difference between the two problems at the boundary of the domain, *i.e.*, for points x with $F(x)$ positive semidefinite but not positive definite. The linear matrix inequality in the SDP given above is equivalent to

$$F(x) \succeq 0, \qquad c \in \operatorname{range}(F(x)), \qquad c^T F(x)^\dagger c \leq t$$

where $F(x)^\dagger$ is the pseudo-inverse of $F(x)$. By eliminating t the SDP is seen to be equivalent to

$$\begin{array}{ll} \text{minimize} & c^T F(x)^\dagger c \\ \text{subject to} & F(x) \succeq 0 \\ & c \in \operatorname{range}(F(x)). \end{array}$$

If $F(x)$ is positive semidefinite but singular, and $c \in \operatorname{range}(F(x))$, the objective function $c^T F(x)^\dagger c$ is finite, whereas $f(x)$ is defined as $+\infty$ in the original problem. However this does not change the optimal value of the problem (unless the set $\{x \mid F(x) \succ 0\}$ is empty).

As an example, consider

$$c = \left[\begin{array}{c} 1 \\ 0 \end{array}\right], \qquad F(x) = \left[\begin{array}{cc} x & 0 \\ 0 & 1-x \end{array}\right].$$

Then the problem in the assignment is to minimize $1/x$, with domain $\{x \mid 0 < x < 1\}$. The optimal value is 1 and is not attained. The SDP reformulation

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \begin{bmatrix} x & 0 & 1 \\ 0 & 1-x & 0 \\ 1 & 0 & t \end{bmatrix} \succeq 0 \end{array}$$

is equivalent to minimizing $1/x$ subject to $0 < x \leq 1$. The optimal value is 1 and attained at $x = 1$.

(d) The cost function can be expressed as

$$f(x) = \bar{c}^T F(x)^{-1} \bar{c} + \mathbf{tr}(F(x)^{-1} S).$$

If we factor S as $S = \sum_{k=1}^m c_k c_k^T$ the problem is equivalent to

$$\text{minimize} \quad \bar{c}^T F(x)^{-1} \bar{c} + \sum_{k=1}^m c_k^T F(x)^{-1} c_k,$$

which we can write as an SDP

$$\begin{array}{ll} \text{minimize} & t_0 + \sum_k t_k \\ \text{subject to} & \begin{bmatrix} F(x) & \bar{c} \\ \bar{c}^T & t_0 \end{bmatrix} \succeq 0 \\ & \begin{bmatrix} F(x) & c_k \\ c_k^T & t_k \end{bmatrix} \succeq 0, \quad k = 1, \ldots, m. \end{array}$$

The variables are t_0, t_1, …, t_m, and x.

7. Exercise A18.4.

Solution. We introduce new variables

$$u = R^2 - r^2, \quad L = \sqrt{w^2 + h^2}$$

and write the problem as the GP

$$\begin{array}{ll} \text{minimize} & 2\pi u L \\ \text{subject to} & (F_1/(2\sigma\pi)) L h^{-1} u^{-1} \leq 1 \\ & (F_2/(2\sigma\pi)) L w^{-1} u^{-1} \leq 1 \\ & (1/w_{\text{max}}) w \leq 1, \quad w_{\text{min}} w^{-1} \leq 1 \\ & (1/h_{\text{max}}) h \leq 1, \quad h_{\text{min}} h^{-1} \leq 1 \\ & 0.21 r^2 u^{-1} \leq 1 \\ & (1/R_{\text{max}}^2) u + (1/R_{\text{max}}^2) r^2 \leq 1 \\ & w^2 L^{-2} + h^2 L^{-2} \leq 1 \end{array}$$

with scalar variables r, w, h, u, L, and implicit constraint that the five variables are positive. The desired values can be recovered from the GP by calculating $R = \sqrt{u + r^2}$.

A geometric programming problem can only have monomial equality constraints, so we cannot add an equality constraint $L^2 = w^2 + h^2$. Therefore we changed it to an inequality

$$L^2 \geq w^2 + h^2,$$

i.e., $w^2 L^{-2} + h^2 L^{-2} \leq 1$. To see why this works, notice that L appears only in the objective and the first two inequality constraints. Each of these involves an expression that is monotonically increasing in L, so at the optimum L will be equal to $\sqrt{w^2 + h^2}$. If this is not the case, then we could make L smaller, which maintains feasibility and strictly decreases the objective.

Also, note that we replaced the inequality $R \geq 1.1r$ with $u \geq (1.1^2 - 1)r^2 = 0.21r^2$.

8. Exercise A4.13.

Solution. We first show that the matrix is nonsingular. Assume

$$\begin{bmatrix} A^{-1} & I \\ B^{-1} & -I \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}.$$

From the first equation, $y = -A^{-1}x$. Substituting this in the second equation gives $B^{-1}x + A^{-1}x = 0$, and therefore $x^T(B^{-1} + A^{-1})x = 0$. Since A and B are positive definite this implies $x = 0$. If $x = 0$, then also $y = -A^{-1}x = 0$. This shows that the columns of the matrix are linearly independent.

Following the hint, we write the constraint as

$$\begin{bmatrix} A^{-1} & B^{-1} \\ I & -I \end{bmatrix} \begin{bmatrix} X & X \\ X & X \end{bmatrix} \begin{bmatrix} A^{-1} & I \\ B^{-1} & -I \end{bmatrix} \preceq \begin{bmatrix} A^{-1} & B^{-1} \\ I & -I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix} \begin{bmatrix} A^{-1} & I \\ B^{-1} & -I \end{bmatrix}.$$

After working out the matrix products we get

$$\begin{bmatrix} (A^{-1} + B^{-1})X(A^{-1} + B^{-1}) & 0 \\ 0 & 0 \end{bmatrix} \preceq \begin{bmatrix} A^{-1} + B^{-1} & 0 \\ 0 & A + B \end{bmatrix}.$$

This shows that the SDP is equivalent to

$$\begin{array}{ll} \text{maximize} & \mathbf{tr}\, X \\ \text{subject to} & X \preceq (A^{-1} + B^{-1})^{-1}. \end{array}$$

We have $\mathbf{tr}\, X \leq \mathbf{tr}((A^{-1} + B^{-1})^{-1})$ for all feasible X, because the trace is the sum of the diagonal elements and the inequality in the constraint implies that $X_{ii} \leq ((A^{-1} + B^{-1})^{-1})_{ii}$ for $i = 1, \ldots, n$. Moreover equality $\mathbf{tr}\, X = \mathbf{tr}((A^{-1} + B^{-1})^{-1})$ holds for $X = (A^{-1} + B^{-1})^{-1}$. This proves that the optimal value is equal to

$$f(A, B) = \mathbf{tr}((A^{-1} + B^{-1})^{-1}).$$

From this we conclude that $f(A,B) = \mathbf{tr}((A^{-1}+B^{-1}))$ is concave. Define a function $F : \mathbf{S}^n \times \mathbf{S}^n \times \mathbf{S}^n \to \mathbf{R}$ with

$$\mathbf{dom}\, F = \{(X,A,B) \mid \left[\begin{array}{cc} X & X \\ X & X \end{array}\right] \preceq \left[\begin{array}{cc} A & 0 \\ 0 & B \end{array}\right]\}.$$

and value $F(X,A,B) = -\mathbf{tr}(X)$ on its domain. This function is convex, jointly in (X,A,B), because its domain is a convex set and on its domain it is linear. Therefore the function $\inf_X F(X,A,B) = -f(A,B)$ is convex.

9. Exercise A4.5.

Solution. As a first solution, one can start by formulating the problem in the assignment as a linear-fractional program

$$\begin{array}{ll} \text{minimize} & s/v \\ \text{subject to} & a_i^T x + b_i \leq s, \quad i=1,\ldots,m \\ & c_i^T x + d_i \geq v, \quad i=1,\ldots,m \\ & Fx \preceq g \end{array}$$

with variables x, s, v, where we define the domain of the objective function to be $\{(s,v) \mid v > 0\}$. We can now use the trick of §4.3.2 and lecture 4, page 21. We make a change of variables

$$y = x/v, \qquad u = s/v, \qquad z = 1/v$$

and obtain an LP

$$\begin{array}{ll} \text{minimize} & u \\ \text{subject to} & a_i^T y + b_i z \leq u, \quad i=1,\ldots,m \\ & c_i^T y + d_i z \geq 1, \quad i=1,\ldots,m \\ & Fy \preceq gz \\ & z \geq 0. \end{array} \tag{4}$$

We can also derive this LP formulation directly. The LP (4) is equivalent to the convex optimization problem

$$\begin{array}{ll} \text{minimize} & \max\limits_{i=1,\ldots,m} (a_i^T y + b_i t) \\ \text{subject to} & \min\limits_{i=1,\ldots,p} (c_i^T y + d_i t) \geq 1 \\ & Fy \preceq gt \\ & t \geq 0 \end{array} \tag{5}$$

with variables y, t. To show that (5) is equivalent to the problem in the assignment, we first note that $t > 0$ for all feasible (y,t) in (5). Indeed, the first constraint implies that $(y,t) \neq 0$. We must have $t > 0$ because otherwise $Fy \preceq 0$ and $y \neq 0$, which means that y defines an unbounded direction in the polyhedron

$\{x \mid Fx \preceq g\}$, contradicting the assumption that this polyhedron is bounded. If $t > 0$ for all feasible y, t, we can rewrite problem (5) as

$$\begin{array}{ll} \text{minimize} & t \max\limits_{i=1,\ldots,m} (a_i^T(y/t) + b_i) \\ \text{subject to} & \min\limits_{i=1,\ldots,p} (c_i^T(y/t) + d_i) \geq 1/t \\ & F(y/t) \preceq g \\ & t \geq 0. \end{array} \tag{6}$$

Next we argue that the first constraint necessarily holds with equality at the optimum, *i.e.*, the optimal solution of (6) is also the solution of

$$\begin{array}{ll} \text{minimize} & t \max\limits_{i=1,\ldots,m} (a_i^T(y/t) + b_i) \\ \text{subject to} & \min\limits_{i=1,\ldots,p} (c_i^T(y/t) + d_i) = 1/t \\ & F(y/t) \preceq g \\ & t \geq 0. \end{array} \tag{7}$$

To see this, suppose we fix y/t in (6) and optimize only over t. Since $\max_i (a_i^T(y/t) + b_i) \geq 0$ if $F(y/t) \leq g$, we minimize the cost function by making t as small as possible, *i.e.*, choosing t such that

$$\min_{i=1,\ldots,p} (c_i^T(y/t) + d_i) = 1/t.$$

The final step is to substitute this expression for the optimal t in the cost function of (7) to get

$$\begin{array}{ll} \text{minimize} & \dfrac{\max\limits_{i=1,\ldots,m} (a_i^T(y/t) + b_i)}{\min\limits_{i=1,\ldots,p} (c_i^T(y/t) + d_i)} \\ \text{subject to} & F(y/t) \preceq g \\ & t \geq 0. \end{array}$$

This is the problem of the assignment with $x = y/t$.

10. Exercise T4.43 (b,c)

Solution.

(b) The inequality $\lambda_1(x) \leq t_1$ holds if and only if $A(x) \preceq t_1 I$, and $\lambda_m(A(x)) \geq t_2$ holds if and only if $A(x) \succeq t_2 I$. Therefore we can minimize $\lambda_1(x) - \lambda_m(x)$ by solving

$$\begin{array}{ll} \text{minimize} & t_1 - t_2 \\ \text{subject to} & t_2 I \preceq A(x) \preceq t_1 I. \end{array}$$

This is an SDP with variables $t_1 \in \mathbf{R}$, $t_2 \in \mathbf{R}$, and $x \in \mathbf{R}^n$.

(c) We first note that the problem is equivalent to

$$\begin{array}{ll} \text{minimize} & \lambda/\gamma \\ \text{subject to} & \gamma I \preceq A(x) \preceq \lambda I \end{array} \tag{8}$$

if we take as domain of the objective $\{(\lambda, \gamma) \mid \gamma > 0\}$. This problem is quasiconvex, and can be solved by bisection. The optimal value is less than or equal to α if and only if the inequalities

$$\lambda \leq \gamma\alpha, \qquad \gamma I \preceq A(x) \preceq \lambda I, \qquad \gamma > 0$$

(with variables γ, λ, x) are feasible.
Following the hint we can also pose the problem as the SDP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & I \preceq sA_0 + y_1 A_1 + \cdots + y_n A_n \preceq tI \\ & s \geq 0 \end{array} \tag{9}$$

with variables t, s, y_1, $\ldots$, y_n. We now verify more carefully that the two problems are equivalent. Let $p^\star$ be the optimal value of (8), and $p^\star_{\text{sdp}}$ the optimal value of the SDP (9).
We first show that $p^\star \geq p^\star_{\text{sdp}}$. Let λ/γ be the objective value of (8), evaluated at a feasible point (γ, λ, x). Define $s = 1/\gamma$, $y = x/\gamma$, $t = \lambda/\gamma$. This yields a feasible point in (9), with objective value $t = \lambda/\gamma$. This proves that $p^\star \geq p^\star_{\text{sdp}}$.
Next, we show that $p^*_{\text{sdp}} \geq p^*$. Suppose that s, y, t are feasible in (9). If $s > 0$, then $\gamma = 1/s$, $x = y/s$, $\lambda = t/s$ are feasible in (8) with objective value t. If $s = 0$, we have

$$I \preceq y_1 A_1 + \cdots + y_n A_n \preceq tI.$$

By assumption there exists a point $\hat{x}$ with $A(\hat{x}) \succ 0$. For $x = \hat{x} + \tau y$, we have

$$A(\hat{x} + \tau y) = A(\hat{x}) + \tau \sum_{i=1}^{n} y_i A_i$$

and

$$A(\hat{x}) + \tau I \preceq A(\hat{x} + \tau y) \preceq A(\hat{x}) + \tau t I.$$

The right-hand side inequality implies that

$$\begin{aligned} \lambda_1(A(\hat{x} + \tau y)) &= \sup_{\|u\|_2 = 1} u^T A(\hat{x} + \tau y) u \\ &\leq \sup_{\|u\|_2 = 1} u^T (A(\hat{x}) + \tau t I) u \\ &= \lambda_1(A(\hat{x})) + \tau t. \end{aligned}$$

Similarly, the left-hand side inequality implies that

$$\begin{aligned} \lambda_m(A(\hat{x} + \tau y)) &= \inf_{\|u\|_2 = 1} u^T A(\hat{x} + \tau y) u \\ &\geq \inf_{\|u\|_2 = 1} u^T (A(\hat{x}) + \tau I) u \\ &= \lambda_m(A(\hat{x})) + \tau. \end{aligned}$$

Therefore,

$$\kappa(A(\hat{x} + \tau y)) = \frac{\lambda_1(A(\hat{x} + \tau y))}{\lambda_m(A(\hat{x} + \tau y)} \leq \frac{\lambda_1(A(\hat{x})) + t\tau}{\lambda_m(A(\hat{x})) + \tau}.$$

Letting τ go to infinity, we can construct feasible points in (8), with objective value arbitrarily close to t. We conclude that if s, y, t are feasible in (9) then $t \geq p^\star$. Minimizing over t yields $p^\star_{\mathrm{sdp}} \geq p^\star$.