Homework 5 solutions

1. Exercise A16.2.

Solution.

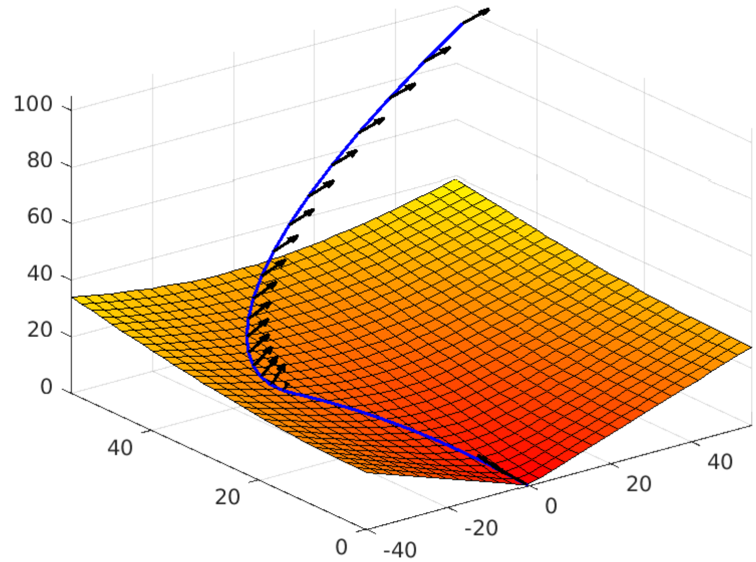
(a) To find the minimum fuel thrust profile for a given K, we solve

$$\begin{array}{ll} \mbox{minimize} & \sum_{k=1}^{K} \|f_k\|_2 \\ \mbox{subject to} & v_{k+1} = v_k + (h/m) f_k - hge_3, \quad k=1,\ldots,K \\ & p_{k+1} = p_k + (h/2)(v_k + v_{k+1}), \quad k=1,\ldots,K \\ & \|f_k\|_2 \leq F^{\max}, \quad k=1,\ldots,K \\ & \alpha\sqrt{(p_k)_1^2 + (p_k)_2^2} \leq (p_k)_3, \quad k=1,\ldots,K \\ & p_{K+1} = 0, \quad v_{K+1} = 0 \\ & p_1 = p(0), \quad v_1 = \dot{p}(0), \end{array}$$

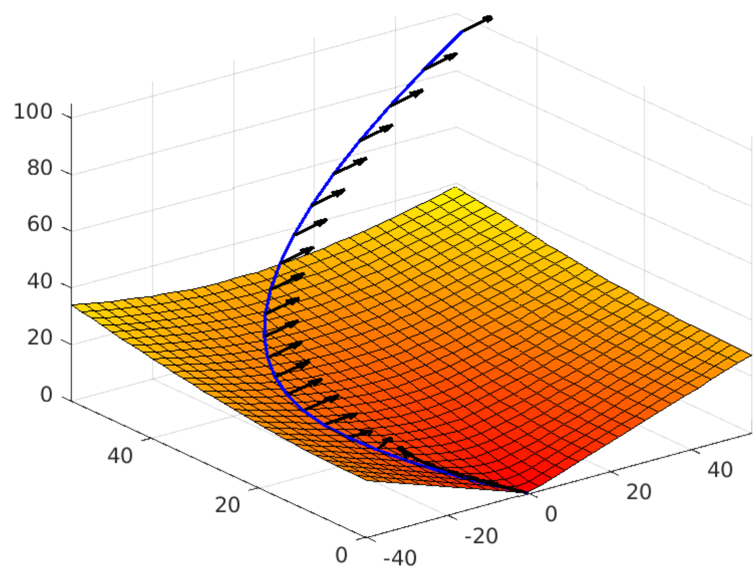
with variables $p_1, \ldots, p_{K+1}$, $v_1, \ldots, v_{K+1}$, and $f_1, \ldots, f_K$. This is a convex optimization problem.

(b) The solution is the smallest K for which the problem in part (a) is feasible. We can find the smallest K by solving the problem in part (a) for a sequence of decreasing values of K until the problem becomes infeasible. We can also use bisection on K.

(c) For part (a) the optimal total fuel consumption is 193.0. For part (b) the minimum touchdown time is $K = 25$. The following plots show the optimal trajectories. The blue line shows the position of the spacecraft, the black arrows show the thrust profile, and the colored surface shows the glide slope constraint.
The first plot shows the minimum fuel trajectory for part (a). Notice that for a portion of the trajectory the thrust is exactly equal to zero.



The second plot is a minimum time trajectory for part (b).



The following MATLAB code was used.

```
spacecraft_landing_data;

% Part (a)
cvx_begin
    variables p(3, K+1) v(3, K+1) f(3, K)
    minimize (sum(norms(f)))
```

```
    subject to
        v(:,2:K+1) == v(:,1:K) + (h/m)*f - h*g*[zeros(2,K); ones(1,K)];
        p(:,2:K+1) == p(:,1:K) + (h/2)*(v(:,1:K) + v(:,2:K+1));
        p(:,1) == p0;
        v(:,1) == v0;
        p(:,K+1) == 0;
        v(:,K+1) == 0;
        p(3, :) >= alpha*norms(p(1:2,:));
        norms(f) <= Fmax;
cvx_end
min_fuel = cvx_optval*gamma*h;
p_minf = p;  v_minf = v;  f_minf = f;

% Part (b); we use a linear search, but bisection is faster
Ki = K;
while(1)
    cvx_begin
        variables p(3,Ki+1) v(3,Ki+1) f(3,Ki)
        minimize(sum(norms(f)))
        subject to
            v(:,2:Ki+1) == v(:,1:Ki) +(h/m)*f - ...
                h*g*[zeros(2,Ki); ones(1,Ki)];
            p(:,2:Ki+1) == p(:,1:Ki) + (h/2)*(v(:,1:Ki)+v(:,2:Ki+1));
            p(:,1) == p0;
            v(:,1) == v0;
            p(:,Ki+1) == 0;
            v(:,Ki+1) == 0;
            p(3,:) >= alpha*norms(p(1:2,:));
            norms(f) <= Fmax;
    cvx_end
    if(strcmp(cvx_status,'Infeasible') == 1)
        Kmin = Ki+1;
        break;
    end
    Ki = Ki-1;
    p_mink = p;  v_mink = v;  f_mink = f;
end

% plot the glide cone
x = linspace(-40,55,30); y = linspace(0,55,30);
[X,Y] = meshgrid(x,y);
Z = alpha*sqrt(X.^2+Y.^2);
figure; colormap autumn; surf(X,Y,Z);
axis([-40,55,0,55,0,105]);
grid on; hold on;
```

```
% plot minimum fuel trajectory for part (a)
plot3(p_minf(1,:), p_minf(2,:), p_minf(3,:), 'b', 'linewidth', 1.5);
quiver3(p_minf(1,1:K), p_minf(2,1:K), p_minf(3,1:K), ...
        f_minf(1,:), f_minf(2,:), f_minf(3,:), 0.3, 'k', 'linewidth',1.5);

% plot minimum time trajectory for part (b)
figure; colormap autumn; surf(X,Y,Z);
axis([-40,55,0,55,0,105]); grid on; hold on;
plot3(p_mink(1,:), p_mink(2,:), p_mink(3,:), 'b', 'linewidth', 1.5);
quiver3(p_mink(1,1:Kmin), p_mink(2,1:Kmin), p_mink(3,1:Kmin),...
        f_mink(1,:), f_mink(2,:), f_mink(3,:), 0.3, 'k', 'linewidth', 1.5);
```

2. Exercise A8.9.

Solution.

(a) The constraint $g(x) \leq \alpha$ is equivalent to

$$\|A_k x + b_k - (c_k^T x + d_k) y^{(k)}\|_2 \leq \alpha(c_k^T x + d_k), \quad k = 1, \ldots, N.$$

This is a set of N convex constraints in x.

(b) The CVX code printed below returns

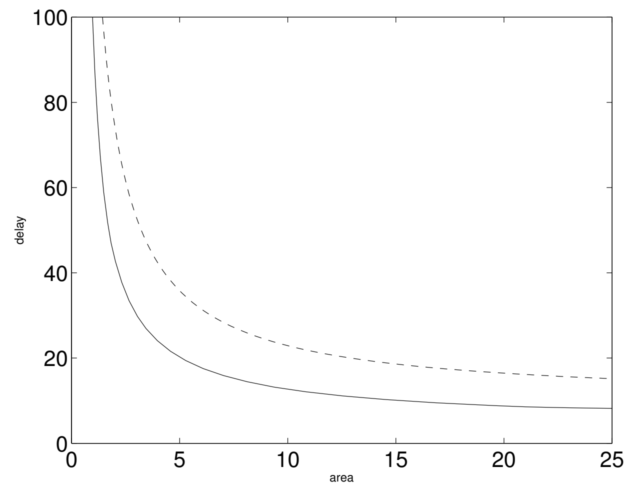
$$x = (4.9, 5.0, 5.2), \qquad t = 0.0495.$$

```
P1 = [1, 0, 0, 0;  0, 1, 0, 0; 0,  0, 1,  0];
P2 = [1, 0, 0, 0;  0, 0, 1, 0; 0, -1, 0, 10];
P3 = [1, 1, 1, -10;  -1, 1, 1, 0; -1, -1, 1, 10];
P4 = [0, 1, 1, 0; 0, -1, 1, 0; -1, 0, 0, 10];
u1 = [0.98; 0.93];
u2 = [1.01; 1.01];
u3 = [0.95; 1.05];
u4 = [2.04; 0.00];
cvx_quiet(true);
l = 0;  u = 1;
tol = 1e-5;
while u-l > tol
    t = (l+u)/2;
    cvx_begin
        variable x(3);
        y1 = P1*[x;1];
        norm( y1(1:2) - y1(3)*u1 ) <= t * y1(3);
        y2 = P2*[x;1];
        norm( y2(1:2) - y2(3)*u2 ) <= t * y2(3);
        y3 = P3*[x;1];
        norm( y3(1:2) - y3(3)*u3 ) <= t * y3(3);
        y4 = P4*[x;1];
```

```
            norm( y4(1:2) - y4(3)*u4 ) <= t * y4(3);
        cvx_end
        disp(cvx_status)
        if cvx_optval == Inf,
            l = t;
        else
            lastx = x;
            u = t;
        end;
    end;
    lastx
```

3. *Interconnect sizing.*

Solution. The dashed line in the figure is the answer for part (a). The solid line is the optimal trade-off curve for part (b).



The figure was generated using the following matlab code.

```
cload1=1.5;
cload2=1;
cload3=5;
maxw=10;
minw=0.1;

N=100;
ws = logspace(-2,1,N);
areas2 = zeros(1,N);
delays2 = zeros(1,N);
for i=1:length(ws)
  w = ws(i)*ones(6,1);
```

```
  areas2(i) = sum(w);
  [T1,T2,T3] = elm_del_example(w);
  delays2(i) = max([T1,T2,T3]);
end;


% set up Mosek input
% variables T, w1, w2, ..., w6

% A1 and C1 give posynomial inequality T1<T
% T1 = (CL1+CL2+CL3)*R1 + CL1*R2 + CL1*R3
%      + C3*R1 + C3*R2 + C3*R3 + C2*R1 + C2*R2
%      + C_1*R1 + C4*R1 + C5*R1 + C6*R1
%    = (CL1+CL2+CL3)/w1 + CL1/w2 + CL1/w3
%      + 3 + w3/w1 + w3/w2 + w2/w1 + w4/w1 + w5/w1 + w6/w1

A1 = [-1 -1  0  0  0  0  0;    % (CL1+CL2+CL3)/(w1*T)
      -1  0 -1  0  0  0  0;    % CL1/(w2*T)
      -1  0  0 -1  0  0  0;    % CL1/(w3*T)
      -1  0  0  0  0  0  0;    % 3/T
      -1 -1  1  0  0  0  0;    % w2/w1
      -1 -1  0  1  0  0  0;    % w3/w1
      -1 -1  0  0  1  0  0;    % w4/w1
      -1 -1  0  0  0  1  0;    % w5/w1
      -1 -1  0  0  0  0  1;    % w6/w1
      -1  0 -1  1  0  0  0 ];  % w3/w2
C1 = [cload1+cload2+cload3; cload1; cload1; 3; ones(6,1)];


% A2 and C2 give posynomial inequality T2<T
% T2 = (CL1+CL2+CL3)*R1 + (CL2+CL3)*R4 + CL2*R5
%      + C5*R1 + C5*R4 + C5*R5 + C4*R1 + C4*R4
%      + C6*R1 + C6*R4 + C1*R1 + C2*R1 + C3*R1
%    = (CL1+CL2+CL3)/w1 + (CL2+CL3)/w4 + CL2/w5
%      + 3 + w5/w1 + w5/w4 + w4/w1 + w6/w1 + w6/w4 + w3/w1

A2 = [-1 -1  0  0  0  0  0;    % (CL1+CL2+CL3)/(w1*T)
      -1  0  0  0 -1  0  0;    % (CL2+CL3)/(w4*T)
      -1  0  0  0  0 -1  0;    % CL2/(w5*T)
      -1  0  0  0  0  0  0;    % 3/T
      -1 -1  1  0  0  0  0;    % w2/w1
      -1 -1  0  1  0  0  0;    % w3/w1
      -1 -1  0  0  1  0  0;    % w4/w1
      -1 -1  0  0  0  1  0;    % w5/w1
      -1 -1  0  0  0  0  1;    % w6/w1
```

```
      -1  0  0  0 -1  1  0;    % w5/w4
      -1  0  0  0 -1  0  1];   % w6/w4
C2 = [cload1+cload2+cload3; cload2+cload3; cload2; 3; ones(7,1)];


% A3 and C3 give posynomial inequality T3<T
% T3 = (CL1+CL2)*R1 + (CL2+CL3)*R4 + CL3*R6
%       + C6*R1 + C6*R4 + C6*R6 + C4*R1 + C4*R4
%       + C1*R1 + C2*R1 + C3*R1 + C5*R1 + C5*R4
%     = (CL1+CL2+CL3)/w1 + (CL2+CL3)/w4 + CL3/w6 + 3
%       + w6/w1 + w6/w4 + w4/w1 + w2/w1 + w3/w1 + w5/w1 + w5/w4

A3 = [-1 -1  0  0  0  0  0;    % (CL1+CL2+CL3)/(w1*T)
      -1  0  0  0 -1  0  0;    % (CL2+CL3)/(w4*T)
      -1  0  0  0  0  0 -1;    % CL3/(w6*T)
      -1  0  0  0  0  0  0;    % 3/T
      -1 -1  1  0  0  0  0;    % w2/w1
      -1 -1  0  1  0  0  0;    % w3/w1
      -1 -1  0  0  1  0  0;    % w4/w1
      -1 -1  0  0  0  1  0;    % w5/w1
      -1 -1  0  0  0  0  1;    % w6/w1
      -1  0  0  0 -1  1  0;    % w5/w4
      -1  0  0  0 -1  0  1];   % w6/w4
C3 = [cload1+cload2+cload3; cload2+cload3; cload3; 3; ones(7,1)];


% A4 and C4 give size constraints minw <= wi <= maxw

A4 = [zeros(6,1) -eye(6);  zeros(6,1) eye(6)];
C4 = [minw*ones(6,1);  ones(6,1)/maxw];


% A0 and C0 give cost function sum wi + mu*T
A0 = eye(7);  C0 = ones(7,1);

A = [A0;A1;A2;A3;A4];
C = [C0;C1;C2;C3;C4];
map = [zeros(size(A0,1),1); ones(size(A1,1),1);
       2*ones(size(A2,1),1); 3*ones(size(A3,1),1);
       3+[1:12]'];

N = 100;
mus = logspace(-3,3,N);
areas = zeros(1,N);
delays = zeros(1,N);
```

```
for i=1:N

   C(1) = mus(i);
   res = mskgpopt(C,A,map);
   x = res.sol.itr.xx;
   delays(i) = exp(x(1));
   areas(i) = sum(exp(x(2:7)));

end;

plot(areas,delays,'-', areas2,delays2,'--');
axis([0 25 0 100]);
xlabel('area');
ylabel('delay');
```

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

Solution.

(b) We first write the problem as

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & f_i(x) \leq t \quad i = 1, \ldots, K \end{array}$$

where $f_i(x) = c_i^T F(x)^{-1} c_i$ with domain $\{x \mid F(x) \succ 0\}$. We then use the Schur complement theorem to write this as an SDP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \left[\begin{array}{cc} F(x) & c_i \\ c_i^T & t \end{array} \right] \succeq 0, \quad i = 1, \ldots, K. \end{array}$$

(c) The cost function can be expressed as

$$f(x) = \lambda_{\max}(F(x)^{-1}),$$

so $f(x) \leq t$ if and only if $F(x)^{-1} \preceq tI$. Using a Schur complement we get

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

with variables x and t. Alternatively, we can minimize the maximum eigenvalue of $F(x)^{-1}$ by maximizing the minimum eigenvalue of $F(x)$. This is an SDP

$$\begin{array}{ll} \text{minimize} & -t \\ \text{subject to} & F(x) \succeq tI \end{array}$$

with variables x and t.

5. *Minimizing symmetric convex functions of eigenvalues.*

Solution.

(a) If $x = 0$, the first constraint reduces to $u = 0$ and the second constraint to $v \geq 0$. Therefore $v = 0$ is optimal.

If $x > 0$, then feasibility in the second inequality requires that $u > 0$ and $v \geq x^2/u$. From the first inequality, $u \leq \sqrt{x}$. Therefore $v \geq x^2/u \geq x^{3/2}$, with equality if $v = x^{3/2}$ and $u = \sqrt{x}$. Therefore $v = x^{3/2}$ is optimal.

Let $X = Q\Lambda Q^T$ be an eigendecomposition of X, with Q orthogonal and Λ diagonal with diagonal elements $\lambda_1, \ldots, \lambda_n$. Define new variables $\tilde{U} = Q^T U \tilde{Q}$ and $\tilde{V} = Q^T \tilde{V} Q$. Since $\mathbf{tr}\, V = \mathbf{tr}\, \tilde{V}$, and

$$\begin{bmatrix} Q^T & 0 \\ 0 & Q^T \end{bmatrix} \begin{bmatrix} X & U \\ U & I \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & Q \end{bmatrix} = \begin{bmatrix} \Lambda & \tilde{U} \\ \tilde{U} & I \end{bmatrix},$$

$$\begin{bmatrix} Q^T & 0 \\ 0 & Q^T \end{bmatrix} \begin{bmatrix} U & X \\ X & V \end{bmatrix} \begin{bmatrix} Q & 0 \\ 0 & Q \end{bmatrix} = \begin{bmatrix} \tilde{U} & \Lambda \\ \Lambda & \tilde{V} \end{bmatrix},$$

the problem is equivalent to

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

with variables $\tilde{U}, \tilde{V} \in \mathbf{S}^m$. At all feasibility points for this SDP,

$$\begin{bmatrix} \lambda_i & \tilde{U}_{ii} \\ \tilde{U}_{ii} & 1 \end{bmatrix} \succeq 0, \qquad \begin{bmatrix} \tilde{U}_{ii} & \lambda_i \\ \lambda_i & \tilde{V}_{ii} \end{bmatrix} \succeq 0, \quad i = 1, \ldots, n,$$

because these 2×2 matrices are principal submatrices of positive semidefinite matrices. By the result for scalar x, these constraints imply that $\tilde{V}_{ii} \geq \lambda_i^{3/2}$. Therefore the optimal value of the SDP is greater than or equal to $\sum_i \lambda_i^{3/2} = f(X)$. Moreover, by taking $\tilde{V} = \Lambda^{3/2}$ and $\tilde{U} = \Lambda^{1/2}$, we obtain a feasible solution with objective value $\mathbf{tr}\, \tilde{V} = \sum_i \lambda_i^{3/2}$, so this is the optimal value of the SDP.

(b) The SDP is equivalent to

$$\begin{array}{ll} \text{minimize} & x^2/(u+1) + u \\ \text{subject to} & u \geq 0, \end{array}$$

with variable u. If $x = 0$ the optimal solution is $u = 0$. If $x > 0$ the cost function has an unconstrained minimizer at $u = |x| - 1 > -1$. If $|x| \geq 1$, the unconstrained minimizer $|x| - 1$ is feasible for the constrained problem, so it is optimal. If $|x| < 1$, the unconstrained minimizer is less than zero, and the cost function is increasing for $u \geq |x| - 1$. In that case $u = 0$ is optimal for the constrained problem. Substituting the optimal u gives

$$\frac{x^2}{u+1} + u = \begin{cases} x^2 & |x| < 1 \\ 2|x| - 1 & |x| \geq 1. \end{cases}$$

The rest of the solution follows the same arguments as in part (a).

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

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

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

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

If $F(x)$ is positive semidefinite but singular, and $c \in \mathrm{range}(F(x))$, 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

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & \left[\begin{array}{ccc} x & 0 & 1 \\ 0 & 1-x & 0 \\ 1 & 0 & t \end{array}\right] \succeq 0 \end{array}$$

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, $\ldots$, 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

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

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

$$\mathbf{dom}\, F = \{(X,A,B) \mid \begin{bmatrix} X & X \\ X & X \end{bmatrix} \preceq \begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}\}.$$

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

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

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

$$\begin{array}{ll} \text{minimize} & \max_{i=1,\ldots,m} (a_i^T y + b_i t) \\ \text{subject to} & \min_{i=1,\ldots,p} (c_i^T y + d_i t) \geq 1 \\ & Fy \preceq gt \\ & t \geq 0 \end{array} \tag{2}$$

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

$$\begin{array}{ll} \text{minimize} & t \max_{i=1,\ldots,m} (a_i^T (y/t) + b_i) \\ \text{subject to} & \min_{i=1,\ldots,p} (c_i^T (y/t) + d_i) \geq 1/t \\ & F(y/t) \preceq g \\ & t \geq 0. \end{array} \tag{3}$$

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

$$\begin{array}{ll} \text{minimize} & t \max_{i=1,\ldots,m} (a_i^T (y/t) + b_i) \\ \text{subject to} & \min_{i=1,\ldots,p} (c_i^T (y/t) + d_i) = 1/t \\ & F(y/t) \preceq g \\ & t \geq 0. \end{array} \tag{4}$$

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

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

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

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

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

$$\begin{array}{ll} \text{minimize} & \lambda/\gamma \\ \text{subject to} & \gamma I \preceq A(x) \preceq \lambda I \end{array} \tag{5}$$

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

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & I \preceq sA_0 + y_1 A_1 + \cdots + y_n A_n \preceq tI \\ & s \geq 0 \end{array} \tag{6}$$

with variables t, s, y_1, $\dots$, 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 $p^\star \geq p^\star_{\rm sdp}$. Let λ/γ be the objective value of (8), evaluated at a feasible point $(\gamma,\ \lambda,\ x)$. 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 $p^\star \geq p^\star_{\rm sdp}$. Next, we show that $p^*_{\rm sdp} \geq p^*$. 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 $p^\star_{\rm sdp} \geq p^\star$.