L. Vandenberghe May 2, 2018

ECE236B homework #4 solutions

1. (a) Suppose the inequality (1) does not hold for some $y \in \mathbf{dom}\, g$. For small positive t,

$$\hat{x} + t(y - x) \in \mathbf{dom}\, f \cap \mathbf{dom}\, g$$

because $\mathbf{dom}\, f$ is an open set and $\mathbf{dom}\, g$ is a convex set. From the Taylor approximation of f at $\hat{x}$, we have

$$f(\hat{x} + t(y - \hat{x})) = f(\hat{x}) + t\nabla f(x)^T(y - \hat{x}) + o(t).$$

Convexity of g implies, by Jensen's inquality, that

$$g(\hat{x} + t(y - \hat{x})) \leq g(\hat{x}) + t(g(y) - g(\hat{x})).$$

Therefore

$$\begin{array}{rl} & f(\hat{x} + t(y - \hat{x})) + g(\hat{x} + t(y - \hat{x})) \\ \leq & f(\hat{x}) + g(\hat{x}) + t(\nabla f(\hat{x})^T(y - \hat{x}) + g(y) - g(\hat{x})) + o(t). \end{array}$$

By assumption, the linear term in t on the right-hand side is negative, so this shows that $\hat{x}$ is not locally optimal.

(b) Suppose f is convex and (1) holds. Consider an arbitrary $y \in \mathbf{dom}\, f \cap \mathbf{dom}\, g$. We have

$$\begin{array}{rcl} f(y) + g(y) & \geq & f(\hat{x}) + \nabla f(\hat{x})^T(y - \hat{x}) + g(y) \\ & \geq & f(\hat{x}) + g(\hat{x}). \end{array}$$

On the first line we use the inequality $f(y) \geq f(\hat{x}) + \nabla f(\hat{x})^T(y - \hat{x})$ for convex differentiable functions. On the second line we use the inequality in (1).

(c) If we define $v = y - \hat{x}$, the optimality condition (1) can be written as

$$\nabla f(\hat{x})^T v + \|\hat{x} + v\|_1 \geq \|\hat{x}\|_1 \quad \text{for all } v.$$

In other words the optimization problem

$$\text{minimize} \quad \nabla f(\hat{x})^T v + \|\hat{x} + v\|_1$$

in the variable v has the solution $v = 0$. The cost function is separable in the elements of v, so the problem is equivalent to n independent problems

$$\text{minimize} \quad \frac{\partial f(\hat{x})}{\partial x_i} v_i + |\hat{x}_i + v_i|$$

for $i = 1, \ldots, n$. If $\hat{x}_i > 0$, then $v_i = 0$ is optimal if and only if $\partial f(\hat{x})/\partial x_i = -1$. If $\hat{x}_i = 0$, then $v_i = 0$ is optimal only if and only if $-1 \leq \partial f(\hat{x})/\partial x_i \leq 1$. If $\hat{x}_i < 0$, then $v_i = 0$ is optimal only if and only if $\partial f(\hat{x})/\partial x_i = 1$.

2. *Exercise T4.8 (e).* Without loss of generality, we assume that $c_1 \leq c_2 \leq \cdots \leq c_n$.

   If α is an integer, the optimal value is

$$c_1 + c_2 + \cdots + c_\alpha,$$

   the sum of the smallest α elements of c. The vector x with $x_1 = \cdots = x_\alpha = 1$, $x_{\alpha+1} = \cdots = x_n = 0$ is optimal.

   Next, suppose α is not an integer, and denote by $\lfloor\alpha\rfloor$ the greatest integer smaller than α. The optimal value is

$$c_1 + c_2 + \cdots + c_{\lfloor\alpha\rfloor} + c_{\lfloor\alpha\rfloor+1}(\alpha - \lfloor\alpha\rfloor).$$

   An optimal choice for x is

$$x_1 = \cdots = x_{\lfloor\alpha\rfloor} = 1, \qquad x_{\lfloor\alpha\rfloor+1} = \alpha - \lfloor\alpha\rfloor, \qquad x_{\lfloor\alpha\rfloor+2} = \cdots = x_n = 0.$$

   In the investment interpretation given in part (d) of the problem, we invest as much as we can in the assets with the highest returns, until the total budget is used.

   In the case of an inequality constraint $\mathbf{1}^T x \leq \alpha$, we can define

$$j = \min\{k = 1, \ldots, n \mid c_k > 0\}$$

   and $j = n+1$ if $c_k \leq 0$ for all k. If $j > \alpha$, we use the same solution as above. If $j \leq \alpha$, we set

$$x_1 = \cdots = x_{j-1} = 1, \qquad x_j = \cdots = x_n = 0.$$

   In the investment interpretation, this corresponds to a situation where we have the option of not investing the total amount. In this case we invest as much as we can in the assets with the highest *nonnegative* returns.

3. *Exercise T4.13.* We consider the constraints separately. The ith constraint is

$$\sum_{j=1}^{n} A_{ij} x_j \leq b_i \quad \text{for all } A_{ij} \in [\bar{A}_{ij} - V_{ij}, A_{ij} \leq \bar{A}_{ij} + V_{ij}].$$

   Given a vector x, the expression on the left-hand side of the inequality is maximized by taking $A_{ij} = \bar{A}_{ij} + V_{ij}$ when $x_j \geq 0$ and $A_{ij} = \bar{A}_{ij} - V_{ij}$ when $x_j > 0$. This shows that the ith constraint is equivalent to a nonlinear convex constraint in x:

$$\sum_{j=1}^{n} \bar{A}_{ij} x_j + V_{ij}|x_j| \leq b_i.$$

   The problem is therefore equivalent to

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{A}x + V|x| \preceq b \end{array} \tag{1}$$

if we define $|x|$ as the vector $|x| = (|x_1|, |x_2|, \ldots, |x_n|)$. This in turn is equivalent to the LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \bar{A}x + Vy \preceq b \\ & -y \preceq x \preceq y \end{array} \tag{2}$$

with variables $x \in \mathbf{R}^n$, $y \in \mathbf{R}^n$.

To see the equivalence between (1) and (2), first note that the constraints $-y \preceq x \preceq y$ mean $y \succeq |x|$. Suppose x is feasible in (1). Then clearly x and $y = |x|$ are feasible in (2). Conversely if x and y are feasible in (2), then x is feasible in (1) because V is component-wise nonnegative and $y \succeq |x|$.

4. *Exercise T4.25.* We first note that the problem is homogeneous in a and b, so we can replace the strict inequalities $a^T x + b > 0$ and $a^T x + b < 0$ with $a^T x + b \geq 1$ and $a^T x + b \leq -1$, respectively. The variables a and b must satisfy

$$\inf_{\|u\|_2 \leq 1} a^T(P_i u + q_i) \geq 1 - b, \quad i = 1, \ldots, K,$$

and

$$\sup_{\|u\|_2 \leq 1} a^T(P_i u + q_i) \leq -1 - b, \quad i = K+1, \ldots, K+L.$$

The left-hand sides can be expressed as

$$\inf_{\|u\|_2 \leq 1} a^T(P_i u + q_i) = -\|P_i^T a\|_2 + q_i^T a$$

and

$$\sup_{\|u\|_2 \leq 1} a^T(P_i u + q_i) = \|P_i^T a\|_2 + q_i^T a.$$

We therefore obtain a set of second-order cone constraints in a, b:

$$\|P_i^T a\|_2 \leq q_i^T a + b - 1, \quad i = 1, \ldots, L, \qquad \|P_i^T a\|_2 \leq -q_i^T a - b - 1, \quad i = K+1, \ldots, K+L.$$

5. *Exercise A7.9.*

   (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
```

6. *Exercise T4.11 (e).* This is equivalent to the LP

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T y + t \\ \text{subject to} & -y \preceq Ax - b \preceq y \\ & -t\mathbf{1} \preceq x \preceq t\mathbf{1}, \end{array}$$

with variables x, y, and t.

To show the equivalence, assume x is given and fixed in the LP, and we minimize over y and t. This problem is easy to solve because it is separable in $y_1, \ldots, y_m, t$, *i.e.*, we can optimize over each of these variables independently. The variable y_k appears in only two constraints, $y_k \leq a_k^T x - b_k \leq y_k$, or, equivalently $y_k \geq |a_k^T x - b_k|$. The optimal choice is $y_k = |a_k^T x - b_k|$. The variable t appears only in the constraints $-t\mathbf{1} \preceq x \preceq t\mathbf{1}$ or, equivalently, $t \geq \max_k |x_k| = \|x\|_\infty$. The optimal choice is $t = \|x\|_\infty$. We conclude that, for any x, the optimal choice of y and t is $y_k = |a_k^T x - b_k|$ and $t = \|x\|_\infty$. Therefore $\mathbf{1}^T y + t = \|Ax - b\|_1 + \|x\|_\infty$ at the optimum. Optimizing jointly over y, t, and x is equivalent to minimizing $\|Ax - b\|_1 + \|x\|_\infty$.

7. *Exercise A3.5.* We describe two solutions for this problem.

   The first solution is to show that the problem is equivalent to the convex optimization problem

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

   with variables y, t. This can be further expressed as an LP by introducing an additional variable u:

$$
\begin{array}{lll}
\text{minimize} & u & \\
\text{subject to} & a_i^T y + b_i t \leq u, & i = 1, \ldots, m \\
& c_i^T y + d_i t \geq 1, & i = 1, \ldots, p \\
& Fy \preceq gt & \\
& t \geq 0. &
\end{array}
$$

   To show that (3) is equivalent to the problem in the assignment, we first note that $t > 0$ for all feasible (y, t). 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 (3) as

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

   Next we argue that the first constraint necessarily holds with equality at the optimum, *i.e.*, the optimal solution of (4) is also the solution of

$$
\begin{array}{ll}
\text{minimize} & t \max_{i=1,\ldots,m} (a_i^T (y/t) + b_i) \\
\text{subject to} & \min_{i=1,\ldots,p} (c_i^T (y/t) + d_i) = 1/t \\
& F(y/t) \preceq g \\
& t \geq 0.
\end{array}
\tag{5}
$$

To see this, suppose we fix y/t in (4) 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 (5) to get

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

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

As an alternative solution, one can start by formulating the problem in the assignment as a linear-fractional program

$$\begin{array}{lll} \text{minimize} & s/v & \\ \text{subject to} & a_i^T x + b_i \leq s, & i = 1, \dots, m \\ & c_i^T x + d_i \geq v, & i = 1, \dots, m \\ & Fx \preceq g & \\ & v \geq 0 & \end{array}$$

with variables x, s, v. We can now use the trick of §4.3.2 and lecture 4, page 20. We make a change of variables

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

and obtain an LP

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

This is the same LP as in the first method.

8. *Exercise T4.27.* To show the equivalence with the problem in the hint, we assume $x \succeq 0$ is fixed, and optimize over v and w. This is a quadratic problem with equality constraints:

$$\begin{array}{ll} \text{minimize} & \begin{bmatrix} v \\ w \end{bmatrix}^T \begin{bmatrix} I & 0 \\ 0 & \mathbf{diag}(x)^{-1} \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} \\ \text{subject to} & \begin{bmatrix} I & B \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = Ax + b. \end{array}$$

The optimality conditions (from lecture 4, page 10) are

$$\begin{bmatrix} v \\ \mathbf{diag}(x)^{-1}w \end{bmatrix} = \begin{bmatrix} I \\ B^T \end{bmatrix} \nu, \qquad \begin{bmatrix} I & B \end{bmatrix} \begin{bmatrix} v \\ w \end{bmatrix} = Ax + b.$$

In other words,

$$v = \nu, \qquad w = \mathbf{diag}(x)B^T\nu, \qquad v + Bw = Ax + b$$

for some ν. Substituting the expressions for v and w in the third equation, we see that ν must satisfy

$$(I + B\,\mathbf{diag}(x)B^T)\nu = Ax + b,$$

and, since the matrix on the left is invertible for $x \succeq 0$,

$$v = \nu = (I+B\,\mathbf{diag}(x)B^T)^{-1}(Ax+b), \qquad w = \mathbf{diag}(x)B^T(I+B\,\mathbf{diag}(x)B^T)^{-1}(Ax+b).$$

Substituting in the objective of the problem in the hint, we obtain

$$v^Tv + w^T\,\mathbf{diag}(x)^{-1}w = (Ax + b)^T(I + B\,\mathbf{diag}(x)B^T)^{-1}(Ax + b).$$

This shows that the problem is equivalent to the problem in the hint.

As in exercise 4.26 we now introduce hyperbolic constraints and formulate the problem in the hint as

$$\begin{array}{ll} \text{minimize} & t + \mathbf{1}^T s \\ \text{subject to} & v^Tv \leq t \\ & w_i^2 \leq s_i x_i, \quad i = 1, \ldots, n \\ & x \succeq 0 \end{array}$$

with variables $t \in \mathbf{R}$, $s, x, w \in \mathbf{R}^n$, $v \in \mathbf{R}^m$. Converting the hyperbolic constraints into second order cone constraints results in an SOCP.