L. Vandenberghe April 18, 2018

ECE236B homework #2 solutions

1. *Exercise T2.12 (d,g).*

   (d) For fixed y, the set $\{x \mid \|x - x_0\|_2 \leq \|x - y\|_2\}$ is a halfspace. This can be seen by squaring the two sides of the inequality $\|x - x_0\|_2 \leq \|x - y\|_2$:

   $$x^T x - 2x_0^T x + x_0^T x_0 \leq x^T x - 2x^T y + y^T y.$$

   The quadratic terms cancel and we get a linear inequality

   $$2(y - x_0)^T x \leq \|y\|_2^2 - \|x_0\|^2.$$

   The set in the problem statement can therefore be expressed as

   $$\bigcap_{y \in S} \{x \mid \|x - x_0\|_2 \leq \|x - y\|_2\},$$

   i.e., an intersection of halfspaces. Hence it is convex.

   (g) The set is convex, in fact a ball.

   $$\begin{aligned} \|x - a\|_2 \leq \theta \|x - b\|_2 &\iff \|x - a\|_2^2 \leq \theta^2 \|x - b\|_2^2 \\ &\iff (1 - \theta^2) x^T x - 2(a - \theta^2 b)^T x + (a^T a - \theta^2 b^T b) \leq 0. \end{aligned}$$

   If $\theta = 1$, this defines a halfspace. If $\theta < 1$, it defines a ball

   $$\{x \mid (x - x_0)^T (x - x_0) \leq R^2\},$$

   with center x_0 and radius R given by

   $$x_0 = \frac{a - \theta^2 b}{1 - \theta^2}, \qquad R = \left(\frac{\theta^2 \|b\|_2^2 - \|a\|_2^2}{1 - \theta^2} + \|x_0\|_2^2 \right)^{1/2}.$$

2. *Polar of a set.*

   (a) The polar is the intersection of halfspaces $\{y \mid y^T x \leq 1\}$, parametrized by $x \in C$, so it is convex.

   (b) If C is a cone, then we have $y^T x \leq 1$ for all $x \in C$ if and only if $y^T x \leq 0$ for all $x \in C$. To see this, suppose that $y^T x > 0$ for some $x \in C$. Then $\alpha x \in C$ for all $\alpha > 0$, so $\alpha y^T x$ can be made arbitrarily large, and in particular, exceeds one for α large enough. Therefore

   $$C^\circ = \{y \mid y^T x \leq 0 \text{ for all } x \in C\}.$$

   This is the negative of the dual cone: $C^\circ = -C^*$.

(c) We note that $y \in C^\circ$ if and only if the optimal value of the optimization problem

$$\begin{array}{ll} \text{maximize} & y^T x \\ \text{subject to} & x \in C \end{array}$$

is less than or equal to one. The following expressions are easy to verify

$$\sup_{x \in C_1} y^T x = \|y\|_2, \qquad \sup_{x \in C_2} y^T x = \|y\|_\infty, \qquad \sup_{x \in C_3} y^T x = \max_k y_k.$$

Therefore

$$C_1^\circ = \{y \mid \|y\|_2 \leq 1\}, \qquad C_2^\circ = \{y \mid \|y\|_\infty \leq 1\}, \qquad C_3^\circ = \{y \mid \max_k y_k \leq 1\}.$$

3. *Exercise A2.10.* The Hessian of f is

$$\nabla^2 f(x) = f(x) \left(qq^T - \mathbf{diag}(\alpha)^{-1} \, \mathbf{diag}(q)^2 \right)$$

where q is the vector $(\alpha_1/x_1, \ldots, \alpha_n/x_n)$. To show that $\nabla^2 f(x)$ is negative semidefinite, we verify that the inequality

$$y^T \nabla^2 f(x) y = f(x) \left((\sum_{k=1}^n \alpha_k y_k / x_k)^2 - \sum_{k=1}^n \alpha_k y_k^2 / x_k^2 \right) \leq 0$$

holds for all y. This follows from the Cauchy-Schwarz inequality $(u^T v)^2 \leq (u^T u)(v^T v)$ applied to the vectors

$$u = (\sqrt{\alpha_1} y_1 / x_1, \ldots, \sqrt{\alpha_n} y_n / x_n), \qquad v = (\sqrt{\alpha_1}, \ldots, \sqrt{\alpha_n}).$$

With this choice of u and v the Cauchy-Schwarz inequality gives

$$\left(\sum_{k=1}^n \alpha_k y_k / x_k \right)^2 \leq \left(\sum_{k=1}^n \alpha_k y_k^2 / x_k^2 \right) \left(\sum_{k=1}^n \alpha_k \right) \leq \left(\sum_{k=1}^n \alpha_k y_k^2 / x_k^2 \right).$$

The second inequality follows from $\sum_{k=1}^n \alpha_k \leq 1$.

4. *Exercise A5.8.*

(a) The objective function is

$$\sum_{k=1}^N (x^T g(t_k) - y_k)^2 = \|Ax - b\|_2^2$$

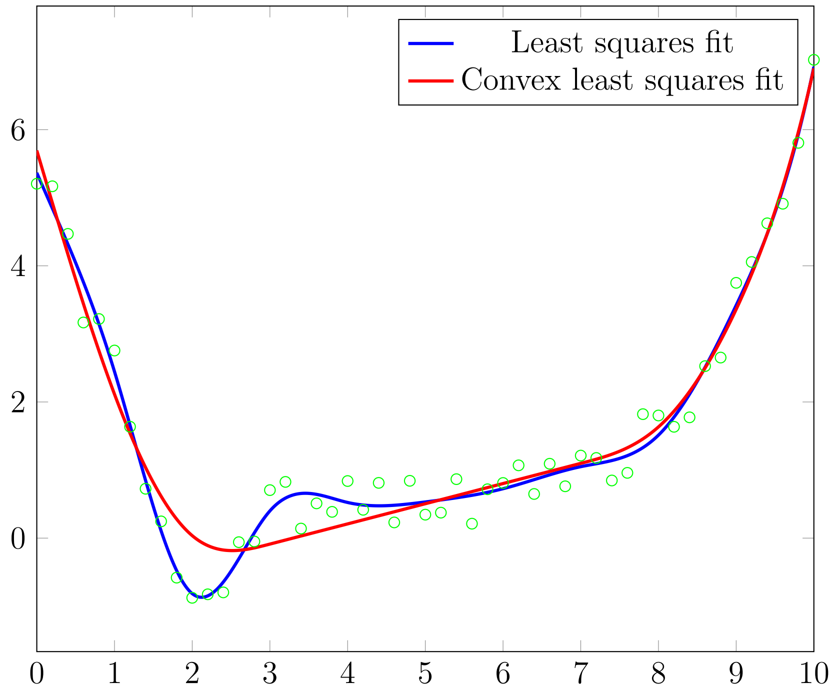
with

$$A = \begin{bmatrix} g(t_1)^T \\ g(t_2)^T \\ \vdots \\ g(t_N)^T \end{bmatrix}, \qquad b = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_N \end{bmatrix}.$$

To handle the convexity constraint we note that f'' is piecewise linear in t. Therefore $f''(t) \geq 0$ for all $t \in (\alpha_0, \alpha_M)$ if and only if $f''(\alpha_k) = x^T g''(\alpha_k) \geq 0$ for $k = 0, \ldots, M$. This gives a set of linear inequalities $Gx \preceq 0$ with

$$G = -\begin{bmatrix} g''(\alpha_0)^T \\ g''(\alpha_1)^T \\ \vdots \\ g''(\alpha_M)^T \end{bmatrix}.$$

(b)
```
[u, y] = spline_data;
N = length(u);
A = zeros(N, 13);
b = y;
for k = 1:N
    [g, gp, gpp] = bsplines(u(k));
    A(k,:) = g';
end;
% Solution without convexity constraint
xls = A\b;
% Solution with convexity constraint
G = zeros(11, 13);
for k = 1:11
    [g, gp, gpp] = bsplines(k-1);
    G(k,:)=  gpp';
end;
cvx_begin
    variable x(13);
    minimize( norm(A*x - b) );
    subject to
        G*x >= 0;
cvx_end
% plot solutions
npts = 1000;
t = linspace(0, 10, npts);
fls = zeros(1, npts);
fcvx = zeros(1, npts);
for k = 1:npts
    [g, gp, gpp] = bsplines(t(k));
    fls(k) = xls' * g;
    fcvx(k) = x' * g;
end;
plot(u, y,'o', t, fls, 'b-', t, fcvx, 'r-');
```



5. *Exercise T2.37.*

(a) It is a closed convex cone, because it is the intersection of (infinitely many) closed halfspaces, and also obviously a cone.

It has nonempty interior because $x = (1, 0, 1, 0 \ldots, 0, 1) \in \textbf{int}\, K_{\rm pol}$. This vector x defines the positive polynomial $p(t) = 1 + t^2 + \cdots + t^{2k}$, and there exists a small positive ϵ such that $y_1 + y_2 t + \cdots + y_{2k+1} t^{2k} \leq p(t)$ for all $\|y\|_2 \leq \epsilon$. Therefore $x - y \in K_{\rm pol}$ for $\|y\|_2 \leq \epsilon$, which shows that $x \in \textbf{int}\, K_{\rm pol}$.

The cone $K_{\rm pol}$ is pointed because $p(t) \geq 0$ and $-p(t) \geq 0$ imply $p(t) = 0$.

(b) Define $v(t) = (1, t, t^2, \ldots, t^k)$. Consider a polynomial $p(t) = x_1 + x_2 t + \cdots + x_{2k+1} t^{2k}$ of degree $2k$ or less and a symmetric matrix $Y \in \mathbf{S}^{2k+1}$. The key observation is that the identity

$$\begin{aligned} p(t) &= v(t)^T Y v(t) \\ &= \begin{bmatrix} 1 \\ t \\ t^2 \\ \vdots \\ t^k \end{bmatrix}^T \begin{bmatrix} Y_{11} & Y_{12} & Y_{13} & \cdots & Y_{1,k+1} \\ Y_{21} & Y_{22} & Y_{23} & \cdots & Y_{2,k+1} \\ Y_{31} & Y_{32} & Y_{33} & \cdots & Y_{3,k+1} \\ \vdots & \vdots & \vdots & & \vdots \\ Y_{k+1,1} & Y_{k+1,2} & Y_{k+1,3} & \cdots & Y_{k+1,k+1} \end{bmatrix} \begin{bmatrix} 1 \\ t \\ t^2 \\ \vdots \\ t^k \end{bmatrix} \end{aligned}$$

holds for all t if and only if if and only if

$$\begin{aligned} x_1 &= Y_{11} \\ x_2 &= Y_{12} + Y_{21} \end{aligned}$$

$$\begin{array}{rcl} x_3 & = & Y_{13}+Y_{22}+Y_{31} \\ & \vdots & \\ x_{2k+1} & = & Y_{k+1,k+1}. \end{array}$$

In other words,

$$x_i = \sum_{m+n=i+1} Y_{mn}, \quad i=1,\ldots,2k+1, \tag{1}$$

where the sum on the right-hand side is over all m and n with $m+n=i+1$. Therefore, if the equations (1) hold and $Y \succeq 0$, then

$$p(t) = v(t)^T Y v(t) \geq 0 \quad \text{for all } t.$$

Conversely, if $p(t) \geq 0$, then we can express the polynomial as $p(t) = r(t)^2 + s(t)^2$ for some polynomials $r(t)$ and $s(t)$ of degree k or less. We write the polynomials as

$$r(t) = a_1 + a_2 t + \cdots + a_{k+1} t^k = a^T v(t)$$

and

$$s(t) = b_1 + b_2 t + \cdots + b_{k+1} t^k = b^T v(t).$$

We have

$$\begin{array}{rcl} p(t) & = & (a^T v(t))^2 + (b^T v(t))^2 \\ & = & v(t)^T a a^T v(t) + v(t)^T b b^T v(t) \\ & = & v(t)^T (a a^T + b b^T) v(t) \\ & = & v(t)^T Y v(t). \end{array}$$

Therefore (1) holds for $Y = aa^T + bb^T$, which is a positive semidefinite matrix.

(c) $z \in K_{\mathrm{pol}}^*$ if and only if $x^T z \geq 0$ for all $x \in K_{\mathrm{pol}}$. Using the previous result, this is equivalent to the condition that

$$\begin{array}{rcl} \displaystyle\sum_{i=1}^{2k+1} z_i \sum_{m+n=i+1} Y_{mn} & = & \displaystyle\sum_{m,n=1}^{k+1} Y_{mn} z_{m+n-1} \\ & = & \mathbf{tr}(Y H(z)) \\ & \geq & 0 \end{array}$$

for all $Y \succeq 0$. This is true if and only if $H(z) \succeq 0$.

(d) The conic hull of the vectors of the form $(1, t, \ldots, t^{2k})$ is the set of nonnegative multiples of all convex combinations of vectors of the form $(1, t, \ldots, t^{2k})$, *i.e.*, nonnegative multiples of vectors of the form $\mathbf{E}(1, t, t^2, \ldots, t^{2k})$.

We have $x \in K_{\mathrm{mom}}^*$ if and only if $x^T y \geq 0$ for all $y \in K_{\mathrm{mom}}$. This is equivalent to the condition that

$$\mathbf{E}(x_1 + x_2 t + x_3 t^2 + \cdots + x_{2k+1} t^{2k}) \geq 0$$

for all distributions on $\mathbf{R}$. This is true if and only if

$$x_1 + x_2 t + x_3 t^2 + \cdots + x_{2k+1} t^{2k} \geq 0$$

for all t.

(e) This follows from the last result in §2.6.1 (K^{**} is the closure of K) and the fact that we have shown that

$$K_{\text{han}} = K_{\text{pol}}^* = K_{\text{mom}}^{**}.$$

For the example, $z \in K_{\text{han}}$ because the matrix

$$H = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is positive semidefinite. On the other hand $z \notin K_{\text{mom}}$ because $\mathbf{E}\, t^2 = 0$ means that the distribution concentrates probability one at $t = 0$. But then we cannot have $\mathbf{E}\, t^4 = 1$.

To construct a sequence of points in K_{mom} that converges to z, consider a discrete distribution with

$$\mathbf{prob}(t = -k) = \frac{1}{2k^4}, \qquad \mathbf{prob}(t = 0) = 1 - \frac{1}{k^4}, \qquad \mathbf{prob}(t = k) = \frac{1}{2k^4}.$$

We have

$$\mathbf{E}(1, t, t^2, t^3, t^4) = (1, 0, \frac{1}{k^2}, 0, 1)$$

and this vector converges to $(1, 0, 0, 0, 1)$ as $k \to \infty$.

6. *Polar of a convex set.*

(a) The property

$$y^T x \leq 1 \ \text{ for all } x \in C,\ y \in C^\circ$$

implies that if $x \in C$, then $y^T x \leq 1$ for all $y \in C^\circ$. In other words $x \in C$ implies that $x \in (C^\circ)^\circ$. This shows that $C \subseteq (C^\circ)^\circ$.

(b) To show that $(C^\circ)^\circ \subseteq C$ we consider a point $x \notin C$ and show that $x \notin (C^\circ)^\circ$.

On page 49, it is shown that if C is a closed convex set and $x \notin C$, then there exist $a \neq 0$ and b such that

$$a^T x > b, \qquad a^T z \leq b \text{ for all } z \in C.$$

Since $0 \in C$, we have $0 \leq b < a^T x$. Define $y = (2/(a^T x + b))a$. Then, for all $z \in C$,

$$y^T z = \frac{2a^T z}{a^T x + b} \leq \frac{2b}{a^T x + b} < 1$$

because $b < a^T x$. Therefore $y \in C^\circ$. On the other hand,

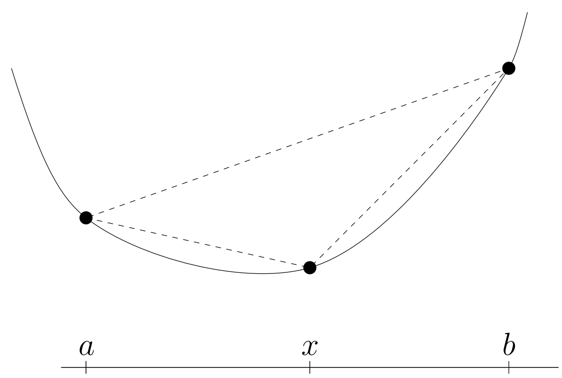
$$y^T x = \frac{2(a^T x)}{a^T x + b} > 1$$

because $a^T x > b$. Therefore $x \notin (C^\circ)^\circ$.

7. *Exercise T3.1.*

   (a) This is Jensen's inequality with $\lambda = (b-x)/(b-a)$.

   (b) We obtain the first inequality by subtracting $f(a)$ from both sides of the inequality in (a). The second inequality follows from subtracting $f(b)$. Geometrically, the inequalities mean that the slope of the line segment between $(a, f(a))$ and $(b, f(b))$ is larger than the slope of the segment between $(a, f(a))$ and $(x, f(x))$, and smaller than the slope of the segment between $(x, f(x))$ and $(b, f(b))$.



   (c) This follows from (b) by taking the limit for $x \to a$ on both sides of the first inequality, and by taking the limit for $x \to b$ on both sides of the second inequality.

   (d) From part (c),

$$\frac{f'(b) - f'(a)}{b - a} \geq 0,$$

   and taking the limit for $b \to a$ shows that $f''(a) \geq 0$.

8. *Exercise T3.18 (a).* Define $g(t) = f(Z + tV)$, where $Z \succ 0$ and $V \in \mathbf{S}^n$, and $\mathbf{dom}\, g = \{t \mid Z + tV \succ 0\}$. We verify that $g(t)$ is a convex function of the scalar variable t.

$$\begin{aligned}
g(t) &= \mathbf{tr}((Z+tV)^{-1}) \\
&= \mathbf{tr}\left(\left(Z^{1/2}(I + tZ^{-1/2}VZ^{-1/2})Z^{1/2}\right)^{-1}\right) \\
&= \mathbf{tr}\left(Z^{-1/2}(I + tZ^{-1/2}VZ^{-1/2})^{-1}Z^{-1/2}\right) \\
&= \mathbf{tr}\left(Z^{-1}(I + tZ^{-1/2}VZ^{-1/2})^{-1}\right) \\
&= \mathbf{tr}\left(Z^{-1}Q(I + t\Lambda)^{-1}Q^T\right)
\end{aligned}$$

$$
\begin{aligned}
&= \mathbf{tr}\left(Q^T Z^{-1} Q (I + t\Lambda)^{-1}\right) \\
&= \sum_{i=1}^{n} (Q^T Z^{-1} Q)_{ii} (1 + t\lambda_i)^{-1}, \qquad (2)
\end{aligned}
$$

where we used the eigenvalue decomposition $Z^{-1/2} V Z^{-1/2} = Q \Lambda Q^T$. The coefficients $(Q^T Z Q)_{ii}$ are positive because they are the diagonal elements of a positive definite matrix $Q^T Z Q$. Also, $\mathbf{dom}\, g = \{t \mid 1 + t\lambda_i > 0,\ i = 1, \ldots, n\}$. The expression (2) shows that g is a positive weighted sum of convex functions $1/(1 + t\lambda_i)$, hence it is convex.