L. Vandenberghe May 30, 2018

ECE236B homework #8 solutions

1. *Exercise A4.15.* This is a convex problem with three equality constraints

$$\begin{array}{ll} \text{minimize} & f_0(X) \\ \text{subject to} & h_1(X) = \alpha \\ & h_2(X) = \beta \\ & h_3(X) = \gamma, \end{array}$$

where $f_0(X) = -\log\det X$ with domain $\mathbf{S}^n_{++}$, and

$$\begin{array}{rcl} h_1(X) & = & \mathbf{tr}\left(\left[\begin{array}{cc} I & 0 \\ 0 & 0 \end{array}\right] X\right), \\ h_2(X) & = & \frac{1}{2}\mathbf{tr}\left(\left[\begin{array}{cc} 0 & I \\ I & 0 \end{array}\right] X\right), \\ h_3(X) & = & \mathbf{tr}\left(\left[\begin{array}{cc} 0 & 0 \\ 0 & I \end{array}\right] X\right). \end{array}$$

The general optimality condition for an equality constrained problem,

$$\nabla f_0(X) + \sum_{i=1}^{3} \nu_i \nabla h_i(X) = 0,$$

reduces to

$$-X^{-1} + \nu_1 \left[\begin{array}{cc} I & 0 \\ 0 & 0 \end{array}\right] + \frac{\nu_2}{2}\left[\begin{array}{cc} 0 & I \\ I & 0 \end{array}\right] + \nu_3 \left[\begin{array}{cc} 0 & 0 \\ 0 & I \end{array}\right] = 0,$$

along with the feasibility conditions $\mathbf{tr}\, X_1 = \alpha$, $\mathbf{tr}\, X_2 = \beta$, $\mathbf{tr}\, X_3 = \gamma$. From the gradient condition

$$X = \left[\begin{array}{cc} \nu_1 I & (\nu_2/2) I \\ (\nu_2/2) I & \nu_3 I \end{array}\right]^{-1} = \left[\begin{array}{cc} \lambda_1 I & \lambda_2 I \\ \lambda_2 I & \lambda_3 I \end{array}\right]$$

where

$$\left[\begin{array}{cc} \lambda_1 & \lambda_2 \\ \lambda_2 & \lambda_3 \end{array}\right] = \left[\begin{array}{cc} \nu_1 & (\nu_2/2) \\ (\nu_2/2) & \nu_3 \end{array}\right]^{-1}.$$

From the feasibility conditions we see that we have to choose λ_i (and hence ν_i), such that

$$\lambda_1 = \alpha/n, \qquad \lambda_2 = \beta/n, \qquad \lambda_3 = \gamma/n.$$

We conclude that the optimal solution is

$$X = \frac{1}{n}\left[\begin{array}{cc} \alpha I & \beta I \\ \beta I & \gamma I \end{array}\right].$$

2. *Exercise A12.12.*

(a) The Lagrangian is

$$\begin{aligned} L(x,Z) &= c^T x + \mathbf{tr}(Z(e_1e_1^T - T_n(x_1,\ldots,x_n))) \\ &= c^Tx + Z_{11} - x_1(Z_{11}+\cdots+Z_{nn}) - 2x_2(Z_{21}+\cdots+Z_{n,n-1}) \\ &\quad - 2x_3(Z_{31}+\cdots+Z_{n,n-2}) - \cdots - 2x_nZ_{n1}. \end{aligned}$$

In the dual SDP we maximize $g(Z) = \inf_x L(x,Z)$ subject to $Z \succeq 0$:

$$\begin{array}{ll} \text{maximize} & Z_{11} \\ \text{subject to} & Z_{11}+Z_{22}+\cdots+Z_{nn} = c_1 \\ & 2(Z_{21}+Z_{32}+\cdots+Z_{n,n-1}) = c_2 \\ & 2(Z_{31}+Z_{42}+\cdots+Z_{n,n-2}) = c_3 \\ & \cdots \\ & 2(Z_{n-1,1}+Z_{n2}) = c_{n-1} \\ & 2Z_{n1} = c_n \\ & Z \succeq 0. \end{array}$$

(b) The constraint $T_n(x_1,\ldots,x_n) \succeq e_1e_1^T$ can be written as

$$\left[\begin{array}{cc} x_1-1 & \bar{x}^T \\ \bar{x} & A \end{array}\right] \succeq 0.$$

where $\bar{x} = (x_2,\ldots,x_n)$ and $A = T_{n-1}(x_1,\ldots,x_{n-1})$. By assumption the 2,2 block A is positive definite, so by the Schur complement theorem the inequality is equivalent to $x_1 - 1 - \bar{x}^TA^{-1}\bar{x} \geq 0$. Hence $x_1 - \bar{x}^TA^{-1}x \geq 1 > 0$ and therefore

$$T_n(x_1,\ldots,x_n) = \left[\begin{array}{cc} x_1 & \bar{x}^T \\ \bar{x} & A \end{array}\right] \succ 0.$$

(c) By strong duality, the primal and dual optimal solutions satisfy $\mathbf{tr}(XZ) = 0$ where

$$X = T_n(x_1,\ldots,x_n) - e_1e_1^T.$$

From part (b), $T_n(x_1,\ldots,x_n)$ is strictly positive definite and therefore the rank of X is at least $n-1$, *i.e.*, its nullspace has dimension at most one. Then $\mathbf{tr}(ZX) = 0$ and $Z \succeq 0$ imply $Z = yy^T$ with y in the nullspace of X. Substituting $Z_{ij} = y_iy_j$ in the equality constraints in the dual SDP gives

$$y_1^2+\cdots+y_n^2 = c_1, \qquad y_1y_k+\cdots+y_{n-k}y_n = c_k/2, \quad k=2,\ldots,n.$$

3. *Exercise A7.1.*

(a) The ellipsoid $\mathcal{E} = \{Q^{1/2}y \mid \|y\|_2 \leq 1\}$ is contained in C if and only if

$$\|Q^{1/2}a_i\|_2 = \sup_{\|y\|_2\leq 1} |a_i^TQ^{1/2}y| \leq 1, \quad i=1,\ldots,p.$$

(b) The dual function is

$$
\begin{aligned}
g(\lambda) &= \inf_{Q \succ 0} L(Q, \lambda) \\
&= \inf_{Q \succ 0} \left(\log\det Q^{-1} + \sum_{i=1}^{p} \lambda_i (a_i^T Q a_i - 1) \right) \\
&= \inf_{Q \succ 0} \left(\log\det Q^{-1} + \mathbf{tr} \left((\sum_{i=1}^{p} \lambda_i a_i a_i^T) Q \right) - \sum_{i=1}^{p} \lambda_i \right).
\end{aligned}
$$

We now use the following fact:

$$
\inf_{X \succ 0} \left(\log\det X^{-1} + \mathbf{tr}(XY) \right) = \begin{cases} \log\det Y + n & Y \succ 0 \\ -\infty & \text{otherwise.} \end{cases}
$$

The value for $Y \succ 0$ follows by setting the gradient of $\log\det X^{-1} + \mathbf{tr}(XY)$ to zero. This gives $-X^{-1} + Y = 0$, so the minimizer is $X = Y^{-1}$ if $Y \succ 0$. If $Y \not\succ 0$, there exists a nonzero a with $a^T Y a \leq 0$. Choosing $X = I + taa^T$ gives $\det X = 1 + t\|a\|_2^2$ and

$$
\log\det X^{-1} + \mathbf{tr}(XY) = -\log(1 + ta^T a) + \mathbf{tr}\, Y + ta^T Y a.
$$

If $a^T Y a \leq 0$ this goes to $-\infty$ as $t \to \infty$.

We conclude that the dual function is

$$
g(\lambda) = \begin{cases} \log\det \sum_{i=1}^{p} (\lambda_i a_i a_i^T) - \sum_{i=1}^{p} \lambda_i + n & \text{if } \sum_{i=1}^{p} (\lambda_i a_i a_i^T) \succ 0 \\ -\infty & \text{otherwise.} \end{cases}
$$

The resulting dual problem is

$$
\begin{array}{ll}
\text{maximize} & \log\det \sum_{i=1}^{p} (\lambda_i a_i a_i^T) - \sum_{i=1}^{p} \lambda_i + n \\
\text{subject to} & \lambda \succeq 0.
\end{array}
$$

(c) The KKT conditions are:

- *Primal feasibility:* $Q \succ 0$ and $a_i^T Q a_i \leq 1$ for $i = 1, \ldots, p$.
- *Nonnegativity of dual multipliers:* $\lambda \succeq 0$.
- *Complementary slackness:* $\lambda_i (1 - a_i^T Q a_i) = 0$ for $i = 1, \ldots, p$.
- *Gradient of Lagrangian is zero:*

$$
Q^{-1} = \sum_{i=1}^{p} \lambda_i a_i a_i^T. \tag{1}
$$

The complementary slackness condition implies that $a_i^T Q a_i = 1$ if $\lambda_i > 0$.

Now suppose Q and λ are primal and dual optimal. If we take the inner product of the two sides of the equation (1) with Q, we get

$$n = \sum_{i=1}^{p} \lambda_i \, \mathbf{tr}(Q a_i a_i^T) = \sum_{i=1}^{p} \lambda_i a_i^T Q a_i = \sum_{i=1}^{p} \lambda_i.$$

The last step follows from the complementary slackness conditions. Finally, we note, again using (1), that

$$x^T Q^{-1} x = \sum_{i=1}^{p} \lambda_i (a_i^T x)^2 \leq \sum_{i=1}^{p} \lambda_i = n$$

if $x \in C$, *i.e.*, if $|a_i^T x| \leq 1$ for $i = 1, \ldots, p$.

4. *Exercise A6.5.*

(a) We write the measurement model as

$$\phi^{-1}(y_i) = a_i^T x + v_i, \quad i = 1, \ldots, m.$$

The function ϕ^{-1} is unknown, but it has derivatives between $1/\beta$ and $1/\alpha$. Therefore $z_i = \phi^{-1}(y_i)$ and y_i must satisfy the inequalities

$$\frac{y_{i+1} - y_i}{\beta} \leq z_{i+1} - z_i \leq \frac{y_{i+1} - y_i}{\alpha}, \qquad i = 1, \ldots, m-1,$$

if we assume that the points are sorted with y_i in increasing order. Conversely, if z and y satisfy these inequalities, then there exists a nonlinear function ϕ with $y_i = \phi(z_i)$, $i = 1, \ldots, m$, and with derivatives between α and β (for example, a piecewise-linear function that interpolates the points). Therefore, as suggested in the problem statement, we can use $z_1, \ldots, z_m$ as parameters instead of ϕ.

The log-likelihood function is

$$l(z, x) = -\frac{1}{2\sigma^2} \sum_{i=1}^{m} (z_i - a_i^T x)^2 - m \log(\sigma\sqrt{2\pi}).$$

Thus to find a maximum likelihood estimate of x and z one solves the problem

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^{m} (z_i - a_i^T x)^2 \\ \text{subject to} & (y_{i+1} - y_i)/\beta \leq z_{i+1} - z_i \leq (y_{i+1} - y_i)/\alpha, \quad i = 1, \ldots, m-1. \end{array}$$

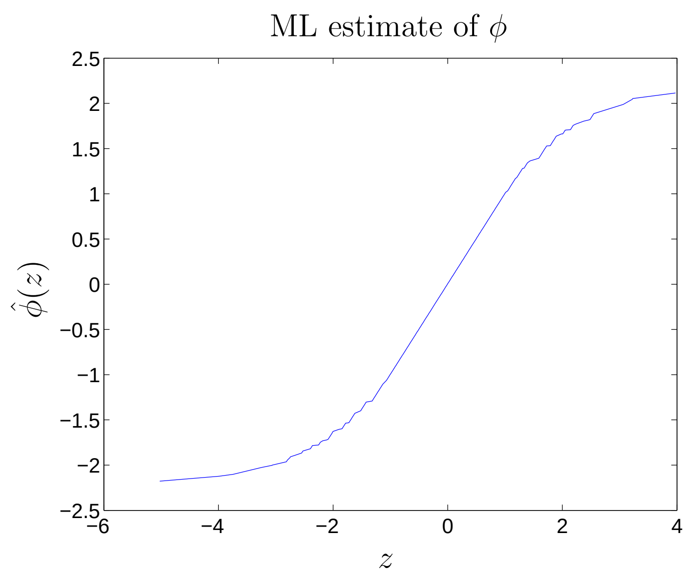
This is a quadratic program with variables $z \in \mathbf{R}^m$ and $x \in \mathbf{R}^n$.

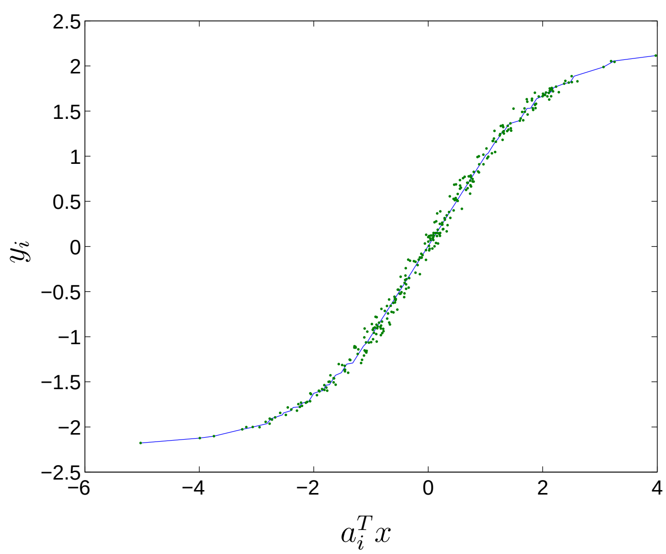
(b) The following MATLAB code solves the problem in the assignment.

```
nonlin_meas_data;
B = [-eye(m-1), zeros(m-1,1)] + [zeros(m-1,1),  eye(m-1)];
cvx_begin
    variables x(n) z(m);
    minimize( norm( z-A*x ) );
    subject to
        (B*y)/beta <= B*z;
        B*z <= (B*y)/alpha;
cvx_end
```

The estimated x is $x = (0.4819, \ -0.4657, \ 0.9364, \ 0.9297)$.

The first figure shows the estimate function ϕ. The second figure shows ϕ and the data points $a_i^T x$, y_i.





5. *Exercise T5.30.* We introduce a Lagrange multiplier $z \in \mathbf{R}^n$ for the equality constraint. The Lagrangian is

$$\begin{aligned} L(X,z) &= \mathbf{tr}\, X - \log\det X + z^T(Xs - y) \\ &= \mathbf{tr}\, X - \log\det X + \mathbf{tr}\left(\frac{1}{2}(zs^T + sz^T)X\right) - y^T s. \end{aligned}$$

On the second line we expressed the linear term $z^T X s$ as an inner product of X with a symmetric matrix $(1/2)(zs^T + sz^T)$. The KKT optimality conditions are:

$$X \succ 0, \qquad Xs = y, \qquad X^{-1} = I + \frac{1}{2}(zs^T + sz^T). \tag{2}$$

(Recall that the gradient of $-\log\det X$ is $-X^{-1}$.) To solve the KKT conditions, we first determine z from the condition $Xs = y$. Multiplying the gradient equation on the right with y gives

$$s = X^{-1}y = y + \frac{1}{2}(z + (z^T y)s). \tag{3}$$

By taking the inner product with y on both sides and simplifying, we get $z^T y = 1 - y^T y$. Substituting in (3) we get

$$z = -2y + (1 + y^T y)s,$$

and substituting this expression for z in (2) gives

$$\begin{aligned} X^{-1} &= I + \frac{1}{2}(-2ys^T - 2sy^T + 2(1 + y^T y)ss^T) \\ &= I + (1 + y^T y)ss^T - ys^T - sy^T. \end{aligned}$$

Finally we verify that this is the inverse of the matrix $X^\star$ given in the problem statement:

$$\begin{aligned} &\left(I + (1 + y^T y)ss^T - ys^T - sy^T\right) X^\star \\ &\quad = (I + yy^T - (1/s^T s)ss^T) + (1 + y^T y)(ss^T + sy^T - ss^T) \\ &\qquad - (ys^T + yy^T - ys^T) - (sy^T + (y^T y)sy^T - (1/s^T s)ss^T) \\ &\quad = I. \end{aligned}$$

To complete the solution, we prove that $X^\star \succ 0$. An easy way to see this is to note that

$$(X^\star)^{-1} = (I - sy^T)(I - ys^T) + ss^T.$$

This matrix is positive semidefinite because

$$v^T (X^\star)^{-1} v = \|(I - ys^T)v\|_2^2 + (s^T v)^2 \geq 0.$$

Moreover at least one of the two terms is strictly positive if $v \neq 0$: if $s^T v = 0$, the first term is $\|v\|_2^2$.

6. *Exercise T5.40.*

$$\begin{array}{ll} \text{minimize} & 1/t \\ \text{subject to} & \sum_{i=1}^p x_i v_i v_i^T \succeq tI \\ & x \succeq 0, \quad \mathbf{1}^T x = 1. \end{array}$$

The Lagrangian is

$$\begin{aligned} L(t,x,Z,z,\nu) &= 1/t - \mathbf{tr}\left(Z(\sum_{i=1}^p x_i v_i v_i^T - tI) \right) - z^T x + \nu(\mathbf{1}^T x - 1) \\ &= 1/t + t\,\mathbf{tr}\, Z + \sum_{i=1}^p x_i(-v_i^T Z v_i - z_i + \nu) - \nu. \end{aligned}$$

The minimum over x_i is bounded below only if $-v_i^T Z v_i - z_i + \nu = 0$. To minimize over t we note that

$$\inf_{t>0}(1/t + t\,\mathbf{tr}\, Z) = \begin{cases} 2\sqrt{\mathbf{tr}\, Z} & \mathbf{tr}\, Z \geq 0 \\ -\infty & \text{otherwise.} \end{cases}$$

The dual function is

$$g(Z,z,\nu) = \begin{cases} 2\sqrt{\mathbf{tr}\, Z} - \nu & v_i^T Z v_i + z_i = \nu, \quad \mathbf{tr}\, Z \geq 0 \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem is

$$\begin{array}{ll} \text{maximize} & 2\sqrt{\mathbf{tr}\, Z} - \nu \\ \text{subject to} & v_i^T Z v_i \leq \nu, \quad i = 1,\ldots,p \\ & Z \succeq 0. \end{array}$$

We can define $W = (1/\nu)Z$,

$$\begin{array}{ll} \text{maximize} & 2\sqrt{\nu}\sqrt{\mathbf{tr}\, W} - \nu \\ \text{subject to} & v_i^T W v_i \geq 1, \quad i = 1,\ldots,p \\ & W \succeq 0. \end{array}$$

Finally, optimizing over ν, gives $\nu = \mathbf{tr}\, W$, so the problem simplifies further to

$$\begin{array}{ll} \text{maximize} & \mathbf{tr}\, W \\ \text{subject to} & v_i^T W v_i \leq 1, \quad i = 1,\ldots,p, \\ & W \succeq 0. \end{array}$$

7. *Exercise A6.1.* With a change of variables

$$z = \frac{1}{\sigma}x, \qquad u = \frac{\mu}{\sigma}, \qquad t = \frac{1}{\sigma},$$

the problem reduces to

$$\text{maximize} \quad m \log t + \sum_{i=1}^m \log f(y_i t - a_i^T z - u),$$

which is a convex optimization problem since the objective is a concave function of (z, u, t). We recover optimal values of x, μ, and σ using

$$\sigma = \frac{1}{t}, \qquad \mu = \frac{u}{t}, \qquad x = \frac{1}{t}z.$$

8. *Exercise A6.9.* The problem is a quasiconvex minimization problem. We can write it as

$$\begin{array}{ll} \text{minimize} & t_1/t_2 \\ \text{subject to} & t_2 I \preceq \sum_{i=1}^m \gamma_i v_i v_i^T \preceq t_1 I \\ & \gamma \succeq 0, \quad \mathbf{1}^T\gamma = 1 \end{array}$$

(where the domain of t_1/t_2 is $\mathbf{R} \times \mathbf{R}_{++}$), with variables γ, t_1, t_2. It can be solved using bisection, solving an SDP feasibility problem in each step.

The problem can also be solved with one SDP, as follows. We divide the linear matrix inequalities by t_2, and define $s = t_1/t_2$, $\lambda = \gamma/t_2$ to get

$$\begin{array}{ll} \text{minimize} & s \\ \text{subject to} & I \preceq \sum_{i=1}^m \lambda_i v_i v_i^T \preceq sI \\ & \lambda \succeq 0, \quad \mathbf{1}^T\gamma = 1/t_2, \end{array}$$

where the variable t_2 must be positive. The last constraint is the same as $\mathbf{1}^T\lambda > 0$, which is redundant. So we end up with the single SDP

$$\begin{array}{ll} \text{minimize} & s \\ \text{subject to} & I \preceq \sum_{i=1}^m \lambda_i v_i v_i^T \preceq sI \\ & \lambda \succeq 0. \end{array}$$

We reconstruct the optimal γ as $\gamma^\star = \lambda^\star/\mathbf{1}^T\lambda^\star$.