L. Vandenberghe May 23, 2018

ECE236B homework #7 solutions

1. *Exercise A4.22.* We make a change of variables $u_i = \log x_i$, $v_j = \log y_j$ and define $\alpha_{ij} = \log A_{ij}$. The GP in convex form is

$$\begin{array}{ll} \text{minimize} & \log\,(\sum_{i=1}^n \sum_{j=1}^n e^{\alpha_{ij}+u_i+v_j}) \\ \text{subject to} & c^T u = 0 \\ & d^T v = 0, \end{array}$$

with variables $u, v \in \mathbf{R}^n$. The optimality conditions are

$$c^T u = d^T v = 0, \qquad \nabla_u L(u,v,\lambda,\gamma) = \nabla_v(u,v,\lambda,\gamma) = 0$$

where L is the Lagrangian

$$L(u,v,\lambda) = \log\,(\sum_{i=1}^n \sum_{j=1}^n e^{\alpha_{ij}+u_i+v_j}) - \lambda c^T u - \gamma d^T v.$$

The optimal u and v therefore satisfy

$$\frac{e^{u_i} \sum_{j=1}^n e^{\alpha_{ij}} e^{v_j}}{\sum_{k=1}^n \sum_{l=1}^n e^{\alpha_{kl}+u_k+v_l}} = \lambda c_i, \quad i = 1,\ldots,n,$$

and

$$\frac{e^{v_j} \sum_{i=1}^n e^{\alpha_{ij}} e^{u_i}}{\sum_{k=1}^n \sum_{l=1}^n e^{\alpha_{kl}+u_k+v_l}} = \gamma d_j, \quad j = 1,\ldots,n$$

for some scalars λ, γ. In the original variables $x_i = e^{u_i}$, $y_i = e^{v_i}$, these equations are

$$\frac{1}{x^T A y}\,\mathbf{diag}(x) A y = \lambda c, \qquad \frac{1}{x^T A y}\,\mathbf{diag}(y) A^T x = \gamma d.$$

Taking the inner product with $\mathbf{1}$ (and using the fact that $\mathbf{1}^T c = \mathbf{1}^T d = 1$) shows $\lambda = \gamma = 1$. Therefore

$$\frac{1}{x^T A y}\,\mathbf{diag}(x) A\,\mathbf{diag}(y)\mathbf{1} = c, \qquad \frac{1}{x^T A y}\,\mathbf{diag}(y) A^T\,\mathbf{diag}(x)\mathbf{1} = d.$$

2. In the primal SDP we must have $x_1 = 0$ at any feasible point. Therefore the optimal value is $p^\star = 0$.

   The Lagrangian is

$$
\begin{aligned}
L(x, Z) &= x_1 - \mathbf{tr}(\begin{bmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{12} & Z_{22} & Z_{23} \\ Z_{13} & Z_{23} & Z_{33} \end{bmatrix} \begin{bmatrix} 0 & x_1 & 0 \\ x_1 & x_2 & 0 \\ 0 & 0 & x_1 + 1 \end{bmatrix}) \\
&= (1 - 2Z_{12} - Z_{33})x_1 - Z_{22}x_2 - Z_{33}.
\end{aligned}
$$

   The infimum over x is minus infinity unless $1 - 2Z_{12} - Z_{33} = Z_{22} = 0$, and equal to $-Z_{33}$ otherwise. Therefore the dual SDP is

$$
\begin{array}{ll}
\text{maximize} & -Z_{33} \\
\text{subject to} & 2Z_{12} + Z_{33} = 1 \\
& Z_{22} = 0 \\
& \begin{bmatrix} Z_{11} & Z_{12} & Z_{13} \\ Z_{12} & Z_{22} & Z_{23} \\ Z_{13} & Z_{23} & Z_{33} \end{bmatrix} \succeq 0.
\end{array}
$$

   Since $Z_{22} = 0$, we have $Z_{12} = Z_{23} = 0$ at any feasible point. The first equality constraint then implies that $Z_{33} = 1$, so the optimal value is $d^\star = -1$.

   We see that there is a finite nonzero gap $p^* - d^* = 1$. We can also note that the primal and dual problems are feasible but not strictly feasible.

3. *Exercise A4.17.*

   (a) Let $A = V\Lambda V^T = \sum_{k=1}^n \lambda_k v_k v_k^T$ be the eigenvalue decomposition of A. If we make a change of variables $Y = V^T XV$ the problem becomes

$$
\begin{array}{ll}
\text{maximize} & \mathbf{tr}(\Lambda Y) \\
\text{subject to} & \mathbf{tr}\, Y = r \\
& 0 \preceq Y \preceq I.
\end{array}
$$

   We can assume that Y is diagonal at the optimum. To see this, note that the objective function and the first constraint only involve the diagonal elements of Y. Moreover, if Y satisfies $0 \preceq Y \preceq I$, then its diagonal elements satisfy $0 \le Y_{ii} \le 1$. Therefore if a non-diagonal Y is optimal, setting its off-diagonal elements to zero yields another feasible matrix with the same objective value.

   To find the optimal value we can therefore solve

$$
\begin{array}{ll}
\text{maximize} & \sum_{i=1}^n \lambda_i Y_{ii} \\
\text{subject to} & \sum_{i=1}^n Y_{ii} = r \\
& 0 \le Y_{ii} \le 1, \quad i = 1, \ldots, n.
\end{array}
$$

Since the eigenvalues λ_i are sorted in nonincreasing order, an optimal solution is $Y_{11} = \cdots = Y_{rr} = 1$, $Y_{r+1,r+1} = \cdots = Y_{nn} = 0$. Converting back to the original variables gives

$$X = \sum_{k=1}^{r} v_k v_k^T.$$

(b) Any function of the form $\sup_{X \in C} \mathbf{tr}(AX)$ is convex in A.

(c) To derive the dual of the problem in part (a), we first write it as a minimization

$$\begin{array}{ll} \text{minimize} & -\mathbf{tr}(AX) \\ \text{subject to} & \mathbf{tr}\, X = r \\ & 0 \preceq X \preceq I. \end{array}$$

The Lagrangian is

$$\begin{aligned} L(X, \nu, U, V) &= -\mathbf{tr}(AX) + \nu(\mathbf{tr}\, X - r) - \mathbf{tr}(UX) + \mathbf{tr}(V(X - I)) \\ &= \mathbf{tr}((-A + \nu I - U + V)X) - r\nu - \mathbf{tr}\, V. \end{aligned}$$

By minimizing over X we obtain the dual function

$$g(\nu, U, V) = \begin{cases} -r\nu - \mathbf{tr}\, V & -A + \nu I - U + V = 0 \\ -\infty & \text{otherwise.} \end{cases}$$

The dual problem is

$$\begin{array}{ll} \text{maximize} & -r\nu - \mathbf{tr}\, V \\ \text{subject to} & A - \nu I = V - U \\ & U \succeq 0, \quad V \succeq 0. \end{array}$$

If we change the dual problem to a minimization and eliminate the variable U, we obtain a dual problem for the SDP in part (a) of the assignment:

$$\begin{array}{ll} \text{minimize} & r\nu + \mathbf{tr}\, V \\ \text{subject to} & A - \nu I \preceq V \\ & V \succeq 0. \end{array}$$

By strong duality, the optimal value of this problem is equal to $f(A)$. We can therefore minimize $f(A(x))$ over x by solving the

$$\begin{array}{ll} \text{minimize} & r\nu + \mathbf{tr}\, V \\ \text{subject to} & A(x) - \nu I \preceq V \\ & V \succeq 0, \end{array}$$

which is an SDP in the variables $\nu \in \mathbf{R}$, $V \in \mathbf{S}^n$, $x \in \mathbf{R}^m$.

4. *Exercise A4.10.*

(a) The Lagrangian is

$$L(x,\nu) = x^T(A^TA + \mathbf{diag}(\nu))x - 2b^TAx + b^Tb - \mathbf{1}^T\nu.$$

The dual function is

$$g(\nu) = \begin{cases} -\mathbf{1}^T\nu - b^TA(A^TA+\mathbf{diag}(\nu))^{\dagger}A^Tb + b^Tb & A^TA+\mathbf{diag}(\nu) \succeq 0, \\ & A^Tb \in \mathcal{R}(A^TA+\mathbf{diag}(\nu)) \\ -\infty & \text{otherwise.} \end{cases}$$

Using Schur complements we can express the dual as an SDP

$$\begin{array}{ll} \text{maximize} & -\mathbf{1}^T\nu - t + b^Tb \\ \text{subject to} & \left[\begin{array}{cc} A^TA+\mathbf{diag}(\nu) & -A^Tb \\ -b^TA & t \end{array}\right] \succeq 0 \end{array}$$

with variables ν and t.

(b) We first write the problem as a minimization problem

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T\nu + t - b^Tb \\ \text{subject to} & \left[\begin{array}{cc} A^TA+\mathbf{diag}(\nu) & -A^Tb \\ -b^TA & t \end{array}\right] \succeq 0. \end{array}$$

We introduce a Lagrange multiplier

$$\left[\begin{array}{cc} Z & z \\ z^T & \lambda \end{array}\right]$$

for the constraint and form the Lagrangian

$$\begin{array}{rcl} \tilde{L}(\nu,t,Z,z,\lambda) & = & \mathbf{1}^T\nu + t - b^Tb - \mathbf{tr}(Z(A^TA+\mathbf{diag}(\nu))) + 2z^TA^Tb - t\lambda \\ & = & (\mathbf{1}-\mathbf{diag}(Z))^T\nu + t(1-\lambda) - b^Tb - \mathbf{tr}(ZA^TA) + 2z^TA^Tb. \end{array}$$

This is unbounded below unless $\mathbf{diag}(Z) = \mathbf{1}$ and $\lambda = 1$. The dual problem of the SDP in part (a) is therefore

$$\begin{array}{ll} \text{minimize} & \mathbf{tr}(A^TAZ) - 2b^TAz + b^Tb \\ \text{subject to} & \mathbf{diag}(Z) = \mathbf{1} \\ & \left[\begin{array}{cc} Z & z \\ z^T & 1 \end{array}\right] \succeq 0. \end{array}$$

To see that this is a relaxation of the original problem we note that the binary least-squares problem is equivalent to

$$\begin{array}{ll} \text{minimize} & \mathbf{tr}(A^TAZ) - 2b^TAz + b^Tb \\ \text{subject to} & \mathbf{diag}(Z) = \mathbf{1} \\ & Z = zz^T. \end{array}$$

In the relaxation we replace $Z = zz^T$ by the weaker constraint $Z \succeq zz^T$, which is equivalent to

$$\begin{bmatrix} Z & z \\ z^T & 1 \end{bmatrix} \succeq 0.$$

(c) We have $\mathbf{E}\, v_k^2 = Z_{kk}$ and

$$\begin{aligned}
\mathbf{E}\, \|Av - b\|_2^2 &= \mathbf{E}(v^T A^T A v - 2b^T A^T v + b^T b) \\
&= \mathbf{tr}(\mathbf{E}(vv^T) A^T A) - 2b^T A^T\, \mathbf{E}\, v + b^T b \\
&= \mathbf{tr}(Z A^T A) - 2b^T A^T z + b^T b.
\end{aligned}$$

(d) The MATLAB code for solving the SDP is as follows.

```
cvx_begin sdp
    variable z(n);
    variable Z(n,n) symmetric;
    minimize( trace(A'*A*Z) - 2*b'*A*z + b'*b )
    subject to
        [Z, z; z' 1] >= 0
        diag(Z)==1;
cvx_end
```

The table lists, for each s, the values of $f(x) = \|Ax - b\|_2$ for $x = \hat{x}$ and the four approximations, and also the lower bound on the optimal value of

$$\begin{array}{ll}
\text{minimize} & \|Ax - b\|_2 \\
\text{subject to} & x_k^2 = 1, \quad k = 1, \ldots, n,
\end{array}$$

obtained by the SDP relaxation.

s	$f(\hat{x})$	$f(x^{(a)})$	$f(x^{(b)})$	$f(x^{(c)})$	$f(x^{(d)})$	lower bound
0.5	4.1623	4.1623	4.1623	4.1623	4.1623	4.0524
1.0	8.3245	12.7299	8.3245	8.3245	8.3245	7.8678
2.0	16.6490	30.1419	16.6490	16.6490	16.6490	15.1804
3.0	24.9735	33.9339	25.9555	25.9555	24.9735	22.1139

- For $s = 0.5$, all heuristics return $\hat{x}$. This is likely to be the global optimum, but that is not necessarily true. However, from the lower bound we know that the global optimum is in the interval $[4.0524, 4.1623]$, so even if 4.1623 is not the global optimum, it is quite close.
- For higher values of s, the result from the first heuristic ($x^{(a)}$) is substantially worse than the SDP heuristics.
- All three SDP heuristics return $\hat{x}$ for $s = 1$ and $s = 2$.
- For $s = 3$, the randomized rounding method returns $\hat{x}$. The other SDP heuristics give slightly higher values.

5. *Exercise A4.8.* The unconstrained problem can be written as an SDP

$$\begin{array}{ll} \text{minimize} & c^T x + t \\ \text{subject to} & F(x) \preceq tI \\ & t \geq 0. \end{array} \tag{1}$$

The dual of this problem is

$$\begin{array}{ll} \text{maximize} & \mathbf{tr}(F_0 Z) \\ \text{subject to} & \mathbf{tr}(F_i Z) + c_i = 0, \quad i = 1, \ldots, m \\ & \mathbf{tr}\, Z + s = 1 \\ & Z \succeq 0, \quad s \geq 0. \end{array} \tag{2}$$

The difference with the original dual problem is the addition of an upper bound $\mathbf{tr}\, Z \leq 1$ (written as $\mathbf{tr}\, Z + s = 1$ for $s \geq 0$). We see that $Z^\star$ satisfies this constraint, with $s > 0$. Therefore it is optimal for (2). By complementary slackness we have $t = 0$ at the optimum of the primal problem (1). The optimal x for (1) is therefore optimal for the original SDP.

6. *Exercise A4.28.* The ith constraint

$$\sup_{a_i \in P_i} \max\{a_i^T x - b_i,\ -a_i^T x + b_i\} = \max\{\sup_{a_i \in P_i} (a_i^T x - b_i),\ \sup_{a_i \in P_i} (-a_i^T x + b_i)\} \leq t_i$$

is satisfied if and only if $\sup_{a_i \in P_i} (a_i^T x - b_i) \leq t_i$ and $\sup_{a_i \in P_i} (-a_i^T x + b_i) \leq t_i$. From duality,

$$\sup_{C_i a_i \preceq d_i} a_i^T x = \inf_{\substack{C_i^T z_i = x \\ z_i \succeq 0}} d_i^T z_i, \qquad \sup_{C_i a_i \preceq d_i} -a_i^T x = \inf_{\substack{C_i^T w_i = -x \\ w_i \succeq 0}} d_i^T w_i.$$

So the i constraint is equivalent to existence of z_i, w_i with

$$d_i^T z_i - b_i \leq t_i, \quad C_i^T z_i = x, \quad z_i \succeq 0, \qquad d_i^T w_i + b_i \leq t_i, \qquad C_i^T w_i = -x, \quad w_i \succeq 0.$$

This results in the QP:

$$\begin{array}{ll} \text{minimize} & \sum_{i=1}^m t_i^2 \\ \text{subject to} & d_i^T z_i - b_i \leq t_i, \quad d_i^T w_i + b_i \leq t_i, \quad i = 1, \ldots, m \\ & x = C_i^T z_i = -C_i^T w_i, \quad i = 1, \ldots, m \\ & z_i \succeq 0, \quad w_i \succeq 0, \quad i = 1, \ldots, m. \end{array}$$

7. *Exercise A4.30.*

   (a) The Lagrangian is

$$L(x, y, z) = c^T x + \frac{1}{\mu} \sum_{i=1}^m \log(1 + e^{\mu y_i}) + z^T (Ax - b - y).$$

The minimum over x is unbounded below unless $A^T z + c = 0$. To find the minimum over y we note the function is separable. Setting the derivative with respect to y_i to zero gives

$$\frac{e^{\mu y_i}}{1 + e^{\mu y_i}} = z_i, \qquad y_i = \frac{1}{\mu} \log \frac{z_i}{1 - z_i}$$

and

$$\begin{aligned} &\inf_{y_i} \left(\frac{1}{\mu} \log(1 + e^{\mu y_i}) - z_i y_i\right) \\ &= \begin{cases} -(1/\mu)(z_i \log z_i + (1 - z_i)\log(1 - z_i)) & 0 \le z_i \le 1 \\ -\infty & \text{otherwise} \end{cases} \end{aligned}$$

(with the interpretation $0 \log 0 = 0$). We therefore obtain the dual

$$\begin{array}{ll} \text{maximize} & -b^T z - \frac{1}{\mu} \sum_{i=1}^{m} (z_i \log z_i + (1 - z_i)\log(1 - z_i)) \\ \text{subject to} & A^T z + c = 0 \\ & 0 \preceq z \preceq \mathbf{1}. \end{array}$$

(b) Plugging in the optimal primal solution $x^\star$ of the LP in (25) gives

$$\begin{aligned} q^\star &\le c^T x^\star + \frac{1}{\mu} \sum_{i=1}^{m} \log(1 + e^{\mu(a_i^T x^\star - b_i)}) \\ &\le p^\star + \frac{m \log 2}{\mu} \end{aligned}$$

because $a_i^T x^\star - b_i \le 0$. Plugging in the optimal dual solution $z^\star$ of the LP in the dual of (25) gives

$$\begin{aligned} q^\star &\ge -b^T z^\star - \frac{1}{\mu} \sum_{i=1}^{m} (z_i^\star \log z_i^\star + (1 - z_i^\star)\log(1 - z_i^\star)) \\ &\ge p^\star \end{aligned}$$

because $u \log u \le 0$ on $[0, 1]$.