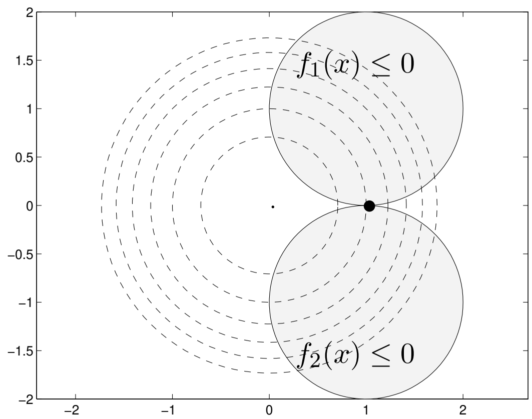
Homework 7 solutions

1. Exercise T5.26.

Solution.

(a) The figure shows the feasible set (the intersection of the two shaded disks) and some contour lines of the objective function. There is only one feasible point, $(1, 0)$, so it is optimal for the primal problem and we have $p^\star = 1$.



(b) The Lagrangian is

$$\begin{aligned} & L(x_1, x_2, \lambda_1, \lambda_2) \\ &= x_1^2 + x_2^2 + \lambda_1((x_1-1)^2 + (x_2-1)^2 - 1) + \lambda_2((x_1-1)^2 + (x_2+1)^2 - 1) \\ &= (1+\lambda_1+\lambda_2)x_1^2 + (1+\lambda_1+\lambda_2)x_2^2 - 2(\lambda_1+\lambda_2)x_1 - 2(\lambda_1-\lambda_2)x_2 + \lambda_1 + \lambda_2. \end{aligned}$$

The KKT conditions are the following.

- x is primal feasible:

$$(x_1-1)^2 + (x_2-1)^2 \leq 1, \qquad (x_1-1)^2 + (x_2+1)^2 \leq 1.$$

- The multipliers for the inequality constraints are nonnegative: $\lambda_1 \geq 0$, $\lambda_2 \geq 0$.
- Complementary slackness:

$$\lambda_1((x_1-1)^2 + (x_2-1)^2 - 1) = \lambda_2((x_1-1)^2 + (x_2+1)^2 - 1) = 0.$$

- The gradient of the Lagrangian at x is zero:

$$\begin{aligned} 2x_1 + 2\lambda_1(x_1-1) + 2\lambda_2(x_1-1) &= 0 \qquad (1) \\ 2x_2 + 2\lambda_1(x_2-1) + 2\lambda_2(x_2+1) &= 0. \end{aligned}$$

At $x = (1, 0)$, the first equation (1) reduces to $2 = 0$, so there exist no λ_1 and λ_2 that satisfy the KKT conditions.

(c) The Lagrange dual function is given by

$$g(\lambda_1, \lambda_2) = \inf_{x_1, x_2} L(x_1, x_2, \lambda_1, \lambda_2).$$

L has a minimum at

$$x_1 = \frac{\lambda_1 + \lambda_2}{1 + \lambda_1 + \lambda_2}, \qquad x_2 = \frac{\lambda_1 - \lambda_2}{1 + \lambda_1 + \lambda_2}$$

if $1 + \lambda_1 + \lambda_2 \geq 0$, and is unbounded below otherwise. Therefore

$$g(\lambda_1, \lambda_2) \quad = \quad \begin{cases} -\dfrac{(\lambda_1 + \lambda_2)^2 + (\lambda_1 - \lambda_2)^2}{1 + \lambda_1 + \lambda_2} + \lambda_1 + \lambda_2 & 1 + \lambda_1 + \lambda_2 \geq 0 \\ -\infty & \text{otherwise,} \end{cases}$$

where we interpret $a/0 = 0$ if $a = 0$ and as $-\infty$ if $a < 0$. The dual problem is

$$\begin{array}{ll} \text{maximize} & \dfrac{\lambda_1 + \lambda_2 - (\lambda_1 - \lambda_2)^2}{1 + \lambda_1 + \lambda_2} \\ \text{subject to} & \lambda_1 \geq 0, \quad \lambda_2 \geq 0. \end{array}$$

Since g is symmetric ($g(\lambda_1, \lambda_2) = g(\lambda_2, \lambda_1)$) and concave, we have

$$\begin{aligned} g(\lambda_1, \lambda_2) &= \frac{1}{2}(g(\lambda_1, \lambda_2) + g(\lambda_2, \lambda_1)) \\ &\leq g(\frac{\lambda_1 + \lambda_2}{2}, \frac{\lambda_1 + \lambda_2}{2}) \end{aligned}$$

for all λ_1 and λ_2. We can therefore take $\lambda_1 = \lambda_2$ in the dual. The dual function

$$g(\lambda_1, \lambda_1) = \frac{2\lambda_1}{1 + 2\lambda_1}$$

tends to the maximum value of 1 as $\lambda_1 = \lambda_2 \to \infty$.
Although we have strong duality ($d^\star = p^\star = 1$), the dual optimum is not attained and therefore the KKT conditions are not solvable.

2. Exercise A5.22.

Solution. We make a change of variables $u_i = \log x_i$, $v_j = \log y_j$ and define $\alpha_{ij} = \log A_{ij}$. The geometric program in convex form is

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

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

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

where L is the Lagrangian

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

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 that $\lambda = \gamma = 1$. Therefore

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

3. Exercise A5.4.

Solution. Define

$$A = \begin{bmatrix} -2y_1^T & 1 \\ -2y_2^T & 1 \\ \vdots & \vdots \\ -2y_5^T & 1 \end{bmatrix}, \qquad b = \begin{bmatrix} d_1^2 - \|y_1\|_2^2 \\ d_2^2 - \|y_2\|_2^2 \\ \vdots \\ d_5^2 - \|y_5\|_2^2 \end{bmatrix}, \qquad C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}, \qquad f = \begin{bmatrix} 0 \\ 0 \\ -1/2 \end{bmatrix},$$

and $z = (x_1, x_2, t)$. With this notation, the problem is

$$\begin{array}{ll} \text{minimize} & \|Az - b\|_2^2 \\ \text{subject to} & z^T C z + 2 f^T z = 0. \end{array}$$

The Lagrangian is

$$L(z,\nu) = z^T(A^T A + \nu C)z - 2(A^T b - \nu f)^T z + \|b\|_2^2,$$

which is bounded below as a function of z only if

$$A^T A + \nu C \succeq 0, \qquad A^T b - \nu f \in \text{range}(A^T A + \nu C).$$

The KKT conditions are as follows.

- *Primal feasibility.*
$$z^T C z + 2f^T z = 0.$$
- *Primal solution minimizes the Lagrangian.* $\inf_z L(z,\nu)$ is finite and z is a minimizer:
$$A^T A + \nu C \succeq 0, \qquad (A^T A + \nu C)z = A^T b - \nu f.$$
Note that the second condition implies the range condition
$$A^T b - \nu f \in \mathrm{range}(A^T A + \nu C).$$

Method 1. We derive the dual problem and solve it via CVX to find the optimal ν. The dual function is

$$g(\nu) = -(A^T b - \nu f)^T (A^T A + \nu C)^\dagger (A^T b - \nu f) + \|b\|_2^2,$$

with domain defined by

$$A^T A + \nu C \succeq 0, \qquad A^T b - \nu f \in \mathrm{range}(A^T A + \nu C).$$

The dual problem can therefore be expressed as an SDP

$$\begin{array}{ll} \text{maximize} & -t \\ \text{subject to} & \begin{bmatrix} A^T A + \nu C & A^T b - \nu f \\ (A^T b - \nu f)^T & t + b^T b \end{bmatrix} \succeq 0. \end{array}$$

Solving this in CVX gives $\nu^\star = 0.5896$. From $\nu^\star$, we get

$$z^\star = (A^T A + \nu C)^{-1}(A^T b - \nu f) = (1.33,\ 0.64,\ 2.18).$$

Hence $x^\star = (1.33, 0.64)$.

Method 2. Alternatively, we can solve the KKT equations directly. To simplify the equations, we make a change of variables

$$w = Q^T L^T z$$

where L is the Cholesky factor in the factorization $A^T A = LL^T$, and Q is the matrix of eigenvectors of $L^{-1} C L^{-T} = Q \Lambda Q^T$. This transforms the KKT equations to

$$w^T \Lambda w + 2g^T w = 0, \qquad I + \nu\Lambda \succeq 0, \qquad (I + \nu\Lambda)w = h - \nu g$$

where

$$g = Q^T L^{-1} f, \qquad h = Q^T L^{-1} A^T b.$$

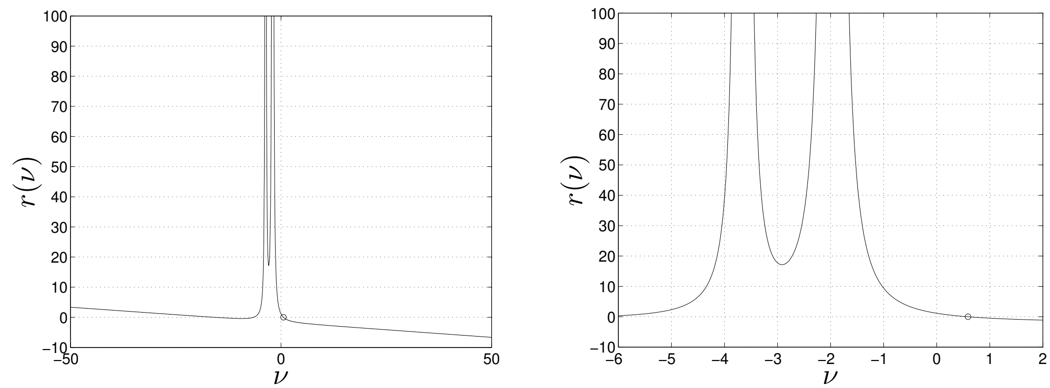
We can eliminate w from the last equation in the KKT conditions to obtain an equation in ν:

$$r(\nu) = \sum_{k=1}^{n+1} \left(\frac{\lambda_k (h_k - \nu g_k)^2}{(1 + \nu\lambda_k)^2} + \frac{2 g_k (h_k - \nu g_k)}{1 + \nu\lambda_k} \right) = 0$$

In our example, the eigenvalues are

$$\lambda_1 = 0.5104, \qquad \lambda_2 = 0.2735, \qquad \lambda_3 = 0.$$

The figure shows the function r on two different scales.

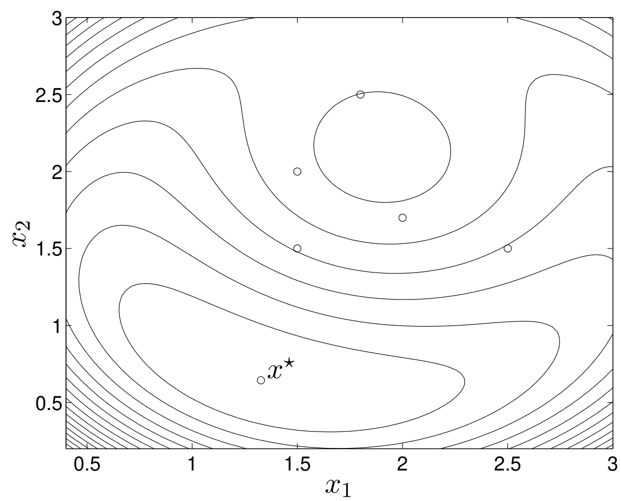


The correct solution of $r(\nu) = 0$ is the one that satisfies $1 + \nu\lambda_k \geq 0$ for $k = 1, 2, 3$, *i.e.*, the solution to the right of the two singularities. This solution can be determined using Newton's method by repeating the iteration

$$\nu := \nu - \frac{r(\nu)}{r'(\nu)}$$

a few times, starting at a value close to the solution. This gives $\nu^\star = 0.5896$. From $\nu^\star$, we determine $x^\star$ as in the first method.

The last figure shows the contour lines and the optimal $x^\star$.



4. Exercise T5.30.

Solution. We introduce a Lagrange multiplier $z \in \mathbf{R}^n$ for the equality constraint. The Lagrangian is

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

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 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 these 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 find that $z^T y = 1 - y^T y$. 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:

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

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 $\|v\|_2^2$.

5. Exercise A15.12.

Solution.

(a) The Lagrangian is

$$
\begin{aligned}
L(x, Z) &= c^T x + \mathbf{tr}(Z(e_1 e_1^T - T_n(x_1, \ldots, x_n))) \\
&= c^T x + Z_{11} - x_1(Z_{11} + \cdots + Z_{nn}) - 2x_2(Z_{21} + \cdots + Z_{n,n-1}) \\
&\quad - 2x_3(Z_{31} + \cdots + Z_{n,n-2}) - \cdots - 2x_n Z_{n1}.
\end{aligned}
$$

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 $T_n(x_1, \ldots, x_n) \succeq e_1 e_1^T$ 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 $A = T_{n-1}(x_1, \ldots, x_{n-1})$. In the induction step we assume that A is positive definite. By the Schur complement theorem this implies that the inequality is equivalent to $x_1 - 1 - \bar{x}^T A^{-1} \bar{x} \geq 0$. Hence $x_1 - \bar{x}^T A^{-1} x \geq 1 > 0$ 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 $XZ = 0$ where

$$
X = T_n(x_1, \ldots, x_n) - e_1 e_1^T
$$

(see lecture 5, page 37). From part (b), $T_n(x_1, \ldots, x_n)$ is strictly positive definite. Therefore the rank of X is at least $n - 1$, *i.e.*, its nullspace has dimension at most one. Then $XZ = 0$ and $Z \succeq 0$ imply $Z = yy^T$ with y in the nullspace of X. Substituting $Z_{ij} = y_i y_j$ in the equality constraints in the dual SDP gives

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

6. Exercise A5.15.

7. Exercise A5.28.

8. *Minimizing sum of largest constraint violations.*

9. Exercise A5.8.