Solution to Practice Problem

EE5138/EE6138

2.1 [L]

Solution. This is readily shown by induction from the definition of convex set. We illustrate the idea for $k = 3$, leaving the general case to the reader. Suppose that $x_1, x_2, x_3 \in C$, and $\theta_1 + \theta_2 + \theta_3 = 1$ with $\theta_1, \theta_2, \theta_3 \geq 0$. We will show that $y = \theta_1 x_1 + \theta_2 x_2 + \theta_3 x_3 \in C$. At least one of the θ_i is not equal to one; without loss of generality we can assume that $\theta_1 \neq 1$. Then we can write

$$y = \theta_1 x_1 + (1 - \theta_1)(\mu_2 x_2 + \mu_3 x_3)$$

where $\mu_2 = \theta_2/(1 - \theta_1)$ and $\mu_2 = \theta_3/(1 - \theta_1)$. Note that $\mu_2, \mu_3 \geq 0$ and

$$\mu_1 + \mu_2 = \frac{\theta_2 + \theta_3}{1 - \theta_1} = \frac{1 - \theta_1}{1 - \theta_1} = 1.$$

Since C is convex and $x_2, x_3 \in C$, we conclude that $\mu_2 x_2 + \mu_3 x_3 \in C$. Since this point and x_1 are in C, $y \in C$.

2.2 [L]

Solution. We prove the first part. The intersection of two convex sets is convex. Therefore if S is a convex set, the intersection of S with a line is convex.

Conversely, suppose the intersection of S with any line is convex. Take any two distinct points x_1 and $x_2 \in S$. The intersection of S with the line through x_1 and x_2 is convex. Therefore convex combinations of x_1 and x_2 belong to the intersection, hence also to S.

2.10 (a) [M]

Solution. A set is convex if and only if its intersection with an arbitrary line $\{\hat{x} + tv \mid t \in \mathbf{R}\}$ is convex.

(a) We have

$$(\hat{x} + tv)^T A(\hat{x} + tv) + b^T(\hat{x} + tv) + c = \alpha t^2 + \beta t + \gamma$$

where

$$\alpha = v^T Av, \qquad \beta = b^T v + 2\hat{x}^T Av, \qquad \gamma = c + b^T\hat{x} + \hat{x}^T A\hat{x}.$$

The intersection of C with the line defined by $\hat{x}$ and v is the set

$$\{\hat{x} + tv \mid \alpha t^2 + \beta t + \gamma \leq 0\},$$

which is convex if $\alpha \geq 0$. This is true for any v, if $v^T A v \geq 0$ for all v, *i.e.*, $A \succeq 0$. The converse does not hold; for example, take $A = -1$, $b = 0$, $c = -1$. Then $A \not\succeq 0$, but $C = \mathbf{R}$ is convex.

2.11 [L]

Solution.

Assume that $\prod_i x_i \geq 1$ and $\prod_i y_i \geq 1$. Using the inequality in the hint, we have

$$\prod_i (\theta x_i + (1-\theta) y_i) \geq \prod x_i^\theta y_i^{1-\theta} = (\prod_i x_i)^\theta (\prod_i y_i)^{1-\theta} \geq 1.$$

2.12 (g) [H]

Solution.

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

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

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

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

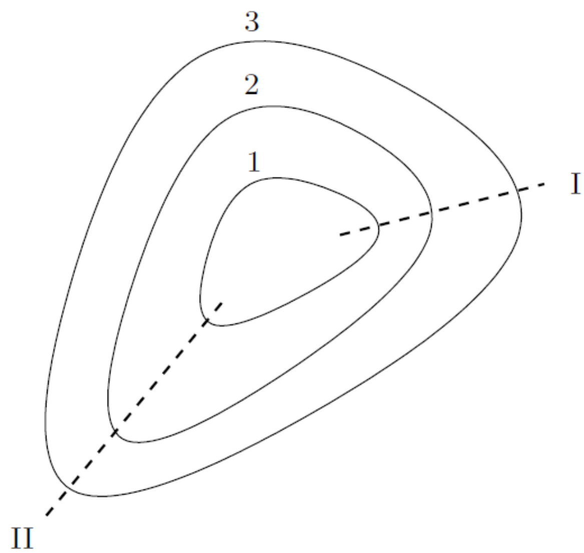
with center x_0 and radius R given by

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

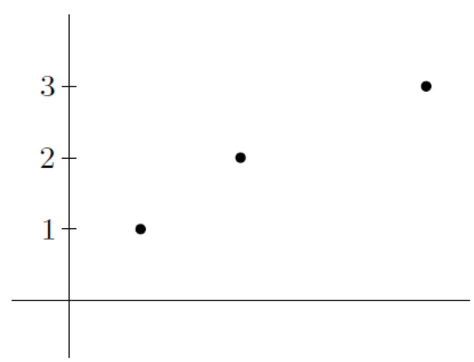
3.2 [L]

Solution. The first function could be quasiconvex because the sublevel sets appear to be convex. It is definitely not concave or quasiconcave because the superlevel sets are not convex.

It is also not convex, for the following reason. We plot the function values along the dashed line labeled I.



Along this line the function passes through the points marked as black dots in the figure below. Clearly along this line segment, the function is not convex.



If we repeat the same analysis for the second function, we see that it could be concave (and therefore it could be quasiconcave). It cannot be convex or quasiconvex, because the sublevel sets are not convex.

3.16 [M]

(a) $f(x) = e^x - 1$ on $\mathbf{R}$.
Solution. Strictly convex, and therefore quasiconvex. Also quasiconcave but not concave.

(b) $f(x_1, x_2) = x_1 x_2$ on $\mathbf{R}^2_{++}$.
Solution. The Hessian of f is

$$\nabla^2 f(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix},$$

which is neither positive semidefinite nor negative semidefinite. Therefore, f is neither convex nor concave. It is quasiconcave, since its superlevel sets

$$\{(x_1, x_2) \in \mathbf{R}^2_{++} \mid x_1 x_2 \geq \alpha\}$$

are convex. It is not quasiconvex.

(c) $f(x_1, x_2) = 1/(x_1 x_2)$ on $\mathbf{R}^2_{++}$.
Solution. The Hessian of f is

$$\nabla^2 f(x) = \frac{1}{x_1 x_2} \begin{bmatrix} 2/(x_1^2) & 1/(x_1 x_2) \\ 1/(x_1 x_2) & 2/x_2^2 \end{bmatrix} \succeq 0$$

Therefore, f is convex and quasiconvex. It is not quasiconcave or concave.

(d) $f(x_1, x_2) = x_1/x_2$ on $\mathbf{R}^2_{++}$.
Solution. The Hessian of f is

$$\nabla^2 f(x) = \begin{bmatrix} 0 & -1/x_2^2 \\ -1/x_2^2 & 2x_1/x_2^3 \end{bmatrix}$$

which is not positive or negative semidefinite. Therefore, f is not convex or concave. It is quasiconvex and quasiconcave (*i.e.*, quasilinear), since the sublevel and superlevel sets are halfspaces.

(e) $f(x_1, x_2) = x_1^2/x_2$ on $\mathbf{R} \times \mathbf{R}_{++}$.
Solution. f is convex, as mentioned on page 72. (See also figure 3.3). This is easily verified by working out the Hessian:

$$\nabla^2 f(x) = \begin{bmatrix} 2/x_2 & -2x_1/x_2^2 \\ -2x_1/x_2^2 & 2x_1^2/x_2^3 \end{bmatrix} = (2/x_2) \begin{bmatrix} 1 \\ -2x_1/x_2 \end{bmatrix} \begin{bmatrix} 1 & -2x_1/x_2 \end{bmatrix} \succeq 0.$$

Therefore, f is convex and quasiconvex. It is not concave or quasiconcave (see the figure).

(f) $f(x_1, x_2) = x_1^\alpha x_2^{1-\alpha}$, where $0 \leq \alpha \leq 1$, on $\mathbf{R}^2_{++}$.
Solution. Concave and quasiconcave. The Hessian is

$$\begin{aligned} \nabla^2 f(x) &= \begin{bmatrix} \alpha(\alpha-1)x_1^{\alpha-2}x_2^{1-\alpha} & \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} \\ \alpha(1-\alpha)x_1^{\alpha-1}x_2^{-\alpha} & (1-\alpha)(-\alpha)x_1^\alpha x_2^{-\alpha-1} \end{bmatrix} \\ &= \alpha(1-\alpha)x_1^\alpha x_2^{1-\alpha} \begin{bmatrix} -1/x_1^2 & 1/x_1x_2 \\ 1/x_1x_2 & -1/x_2^2 \end{bmatrix} \\ &= -\alpha(1-\alpha)x_1^\alpha x_2^{1-\alpha} \begin{bmatrix} 1/x_1 \\ -1/x_2 \end{bmatrix} \begin{bmatrix} 1/x_1 \\ -1/x_2 \end{bmatrix}^T \\ &\preceq 0. \end{aligned}$$

f is not convex or quasiconvex.

3.17 [H]

Suppose $p < 1$, $p \neq 0$. Show that the function

$$f(x) = \left(\sum_{i=1}^{n} x_i^p\right)^{1/p}$$

with $\mathbf{dom}\, f = \mathbf{R}^n_{++}$ is concave. This includes as special cases $f(x) = (\sum_{i=1}^n x_i^{1/2})^2$ and the *harmonic mean* $f(x) = (\sum_{i=1}^n 1/x_i)^{-1}$. *Hint.* Adapt the proofs for the log-sum-exp function and the geometric mean in §3.1.5.

Solution. The first derivatives of f are given by

$$\frac{\partial f(x)}{\partial x_i} = (\sum_{i=1}^{n} x_i^p)^{(1-p)/p} x_i^{p-1} = \left(\frac{f(x)}{x_i}\right)^{1-p}.$$

The second derivatives are

$$\frac{\partial^2 f(x)}{\partial x_i \partial x_j} = \frac{1-p}{x_i}\left(\frac{f(x)}{x_i}\right)^{-p}\left(\frac{f(x)}{x_j}\right)^{1-p} = \frac{1-p}{f(x)}\left(\frac{f(x)^2}{x_i x_j}\right)^{1-p}$$

for $i \neq j$, and

$$\frac{\partial^2 f(x)}{\partial x_i^2} = \frac{1-p}{f(x)}\left(\frac{f(x)^2}{x_i^2}\right)^{1-p} - \frac{1-p}{x_i}\left(\frac{f(x)}{x_i}\right)^{1-p}.$$

We need to show that

$$y^T \nabla^2 f(x) y = \frac{1-p}{f(x)}\left(\left(\sum_{i=1}^{n} \frac{y_i f(x)^{1-p}}{x_i^{1-p}}\right)^2 - \sum_{i=1}^{n} \frac{y_i^2 f(x)^{2-p}}{x_i^{2-p}}\right) \leq 0$$

This follows by applying the Cauchy-Schwarz inequality $a^T b \leq \|a\|_2 \|b\|_2$ with

$$a_i = \left(\frac{f(x)}{x_i}\right)^{-p/2}, \qquad b_i = y_i \left(\frac{f(x)}{x_i}\right)^{1-p/2},$$

and noting that $\sum_i a_i^2 = 1$.

3.22 (a) & (b) [M]

(a) $f(x) = -\log(-\log(\sum_{i=1}^m e^{a_i^T x+b_i}))$ on $\mathbf{dom}\, f = \{x \mid \sum_{i=1}^m e^{a_i^T x+b_i} < 1\}$. You can use the fact that $\log(\sum_{i=1}^n e^{y_i})$ is convex.

Solution. $g(x) = \log(\sum_{i=1}^m e^{a_i^T x+b_i})$ is convex (composition of the log-sum-exp function and an affine mapping), so $-g$ is concave. The function $h(y) = -\log y$ is convex and decreasing. Therefore $f(x) = h(-g(x))$ is convex.

(b) $f(x,u,v) = -\sqrt{uv - x^T x}$ on $\mathbf{dom}\, f = \{(x,u,v) \mid uv > x^T x,\ u,\ v > 0\}$. Use the fact that $x^T x/u$ is convex in (x,u) for $u > 0$, and that $-\sqrt{x_1 x_2}$ is convex on $\mathbf{R}_{++}^2$.

Solution. We can express f as $f(x,u,v) = -\sqrt{u(v - x^T x/u)}$. The function $h(x_1,x_2) = -\sqrt{x_1 x_2}$ is convex on $\mathbf{R}_{++}^2$, and decreasing in each argument. The functions $g_1(u,v,x) = u$ and $g_2(u,v,x) = v - x^T x/u$ are concave. Therefore $f(u,v,x) = h(g(u,v,x))$ is convex.

4.11 (a) and (b) [L]

Solution.

(a) Equivalent to the LP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & Ax - b \preceq t\mathbf{1} \\ & Ax - b \succeq -t\mathbf{1}. \end{array}$$

in the variables x, t. To see the equivalence, assume x is fixed in this problem, and we optimize only over t. The constraints say that

$$-t \le a_k^T x - b_k \le t$$

for each k, *i.e.*, $t \ge |a_k^T x - b_k|$, *i.e.*,

$$t \ge \max_k |a_k^T x - b_k| = \|Ax - b\|_\infty.$$

Clearly, if x is fixed, the optimal value of the LP is $p^\star(x) = \|Ax - b\|_\infty$. Therefore optimizing over t and x simultaneously is equivalent to the original problem.

(b) Equivalent to the LP

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T s \\ \text{subject to} & Ax - b \preceq s \\ & Ax - b \succeq -s. \end{array}$$

Assume x is fixed in this problem, and we optimize only over s. The constraints say that

$$-s_k \le a_k^T x - b_k \le s_k$$

for each k, *i.e.*, $s_k \ge |a_k^T x - b_k|$. The objective function of the LP is separable, so we achieve the optimum over s by choosing

$$s_k = |a_k^T x - b_k|,$$

and obtain the optimal value $p^\star(x) = \|Ax - b\|_1$. Therefore optimizing over t and s simultaneously is equivalent to the original problem.

4.15 [L]

Solution.

(a) The feasible set of the relaxation includes the feasible set of the Boolean LP. It follows that the Boolean LP is infeasible if the relaxation is infeasible, and that the optimal value of the relaxation is less than or equal to the optimal value of the Boolean LP.

(b) The optimal solution of the relaxation is also optimal for the Boolean LP.

4.20 [L]

Solution.

$$\begin{array}{ll} \text{minimize} & \max_i(\sum_{k\neq i} G_{ik}p_k + \sigma_i)/(G_{ii}p_i) \\ & 0 \leq p_i \leq P_i^{\max} \\ & \sum_{k\in K_l} p_k \leq P_l^{\text{gp}} \\ & \sum_{k=1}^{n} G_{ik}p_k \leq P_i^{\text{rc}} \end{array}$$

4.24 (the case of p=2) [M]

Solution.

(a) Minimizing $\|Ax-b\|_2$ is equivalent to minimizing its square. So, let us expand $\|Ax-b\|_2^2$ around the real and complex parts of $Ax-b$:

$$\begin{aligned} \|Ax-b\|_2^2 &= \|\Re(Ax-b)\|_2^2 + \|\Im(Ax-b)\|_2^2 \\ &= \|\Re A \Re x - \Im A \Im x - \Re b\|_2^2 + \|\Re A \Im x + \Im A \Re x - \Im b\|_2^2. \end{aligned}$$

If we define $z^T = [\Re x^T \ \Im x^T]$ as requested, then this becomes

$$\begin{aligned} \|Ax-b\|_2^2 &= \|[\Re A \ \ -\Im A]z - \Re b\|^2 + \|[\Im A \ \Re A]z - \Im b\|^2 \\ &= \left\| \begin{bmatrix} \Re A & -\Im A \\ \Im A & \Re A \end{bmatrix} z - \begin{bmatrix} \Re b \\ \Im b \end{bmatrix} \right\|_2^2. \end{aligned}$$

The values of F and g can be extracted from the above expression.

4.26 (a) [M]

Solution.

(a) The problem is equivalent to

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T t \\ \text{subject to} & t_i(a_i^T x - b_i) \geq 1, \quad i = 1,\dots,m \\ & t \succeq 0. \end{array}$$

Writing the hyperbolic constraints as SOC constraints yields an SOCP

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T t \\ \text{subject to} & \left\| \begin{bmatrix} 2 \\ a_i^T x - b_i - t_i \end{bmatrix} \right\|_2 \leq a_i^T x + b_i + t_i, \quad i = 1,\dots,m \\ & t_i \geq 0, \quad a_i^T x - b_i \geq 0, \quad i = 1,\dots,m. \end{array}$$

4.33 [M]

Solution.

(a) This is equivalent to the GP

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & p(x)/t \le 1, \quad q(x)/t \le 1. \end{array}$$

Now make the logarithmic change of variables $x_i = e^{y_i}$.

(b) Equivalent to

$$\begin{array}{ll} \text{minimize} & \exp(t_1) + \exp(t_2) \\ \text{subject to} & p(x) \le t_1, \quad q(x) \le t_2. \end{array}$$

Now make the logarithmic change of variables $x_i = e^{y_i}$ (but not to t_1, t_2).

(c) Equivalent to

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & p(x) \le t(r(x) - q(x)), \end{array}$$

and

$$\begin{array}{ll} \text{minimize} & t \\ \text{subject to} & (p(x)/t + q(x))/r(x) \le 1, \end{array}$$

which is a GP.

4.40 (b) (for the two cases of QP and QCQP) [H]

Solution.

(a) QP. Express $P = WW^T$ with $W \in \mathbf{R}^{n \times r}$.

$$\begin{array}{ll} \text{minimize} & t + 2q^T x + 2r \\ \text{subject to} & \begin{bmatrix} I & W^T x \\ x^T W & tI \end{bmatrix} \succeq 0 \\ & \mathbf{diag}(Gx - h) \preceq 0 \\ & Ax = b, \end{array}$$

with variables x, $t \in \mathbf{R}$.

(b) QCQP. Express $P_i = W_i W_i^T$ with $W_i \in \mathbf{R}^{n \times r_i}$.

$$\begin{array}{ll} \text{minimize} & t_0 + 2q_0^T x + 2r_0 \\ \text{subject to} & t_i + 2q_i^T x + 2r_i < 0. \quad i = 1, \ldots, m \\ & \begin{bmatrix} I & W_i^T x \\ x^T W_i & t_i I \end{bmatrix} \succeq 0, \quad i = 0, 1, \ldots, m \\ & Ax = b, \end{array}$$

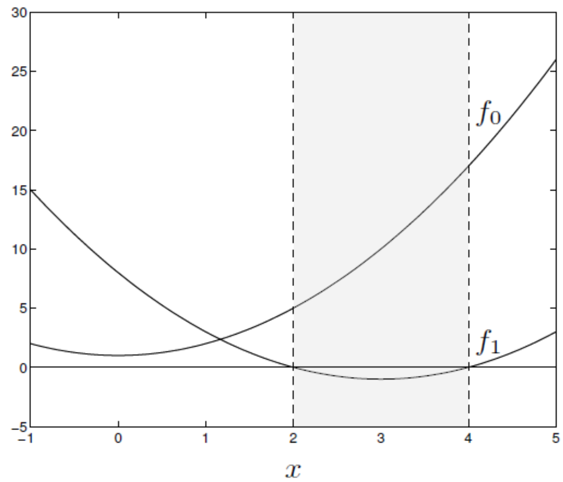
with variables x, $t_i \in \mathbf{R}$.

5.1(a)-(c) [L]

Solution.

(a) The feasible set is the interval $[2, 4]$. The (unique) optimal point is $x^\star = 2$, and the optimal value is $p^\star = 5$.

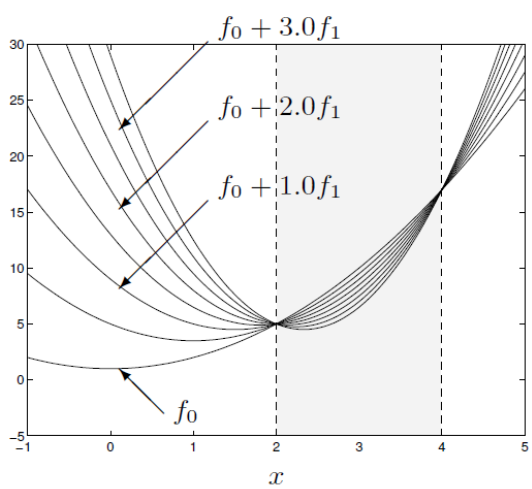
The plot shows f_0 and f_1.



(b) The Lagrangian is

$$L(x, \lambda) = (1 + \lambda)x^2 - 6\lambda x + (1 + 8\lambda).$$

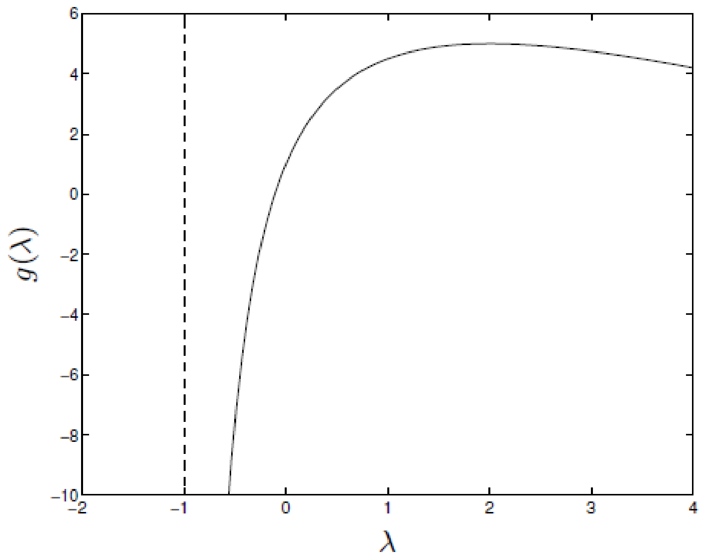
The plot shows the Lagrangian $L(x, \lambda) = f_0 + \lambda f_1$ as a function of x for different values of $\lambda \geq 0$. Note that the minimum value of $L(x, \lambda)$ over x (*i.e.*, $g(\lambda)$) is always less than $p^\star$. It increases as λ varies from 0 toward 2, reaches its maximum at $\lambda = 2$, and then decreases again as λ increases above 2. We have equality $p^\star = g(\lambda)$ for $\lambda = 2$.



For $\lambda > -1$, the Lagrangian reaches its minimum at $\tilde{x} = 3\lambda/(1+\lambda)$. For $\lambda \leq -1$ it is unbounded below. Thus

$$g(\lambda) = \begin{cases} -9\lambda^2/(1+\lambda) + 1 + 8\lambda & \lambda > -1 \\ -\infty & \lambda \leq -1 \end{cases}$$

which is plotted below.



We can verify that the dual function is concave, that its value is equal to $p^\star = 5$ for $\lambda = 2$, and less than $p^\star$ for other values of λ.

(c) The Lagrange dual problem is

$$\begin{array}{ll} \text{maximize} & -9\lambda^2/(1+\lambda) + 1 + 8\lambda \\ \text{subject to} & \lambda \geq 0. \end{array}$$

The dual optimum occurs at $\lambda = 2$, with $d^\star = 5$. So for this example we can directly observe that strong duality holds (as it must — Slater's constraint qualification is satisfied).

5.11 [H]

Solution. The Lagrangian is

$$L(x, z_1, \ldots, z_N) = \sum_{i=1}^{N} \|y_i\|_2 + \frac{1}{2}\|x - x_0\|_2^2 + \sum_{i=1}^{N} z_i^T(y_i - A_i x - b_i).$$

We first minimize over y_i. We have

$$\inf_{y_i}(\|y_i\|_2 + z_i^T y_i) = \begin{cases} 0 & \|z_i\|_2 \leq 1 \\ -\infty & \text{otherwise.} \end{cases}$$

(If $\|z_i\|_2 > 1$, choose $y_i = -tz_i$ and let $t \to \infty$, to show that the function is unbounded below. If $\|z_i\|_2 \leq 1$, it follows from the Cauchy-Schwarz inequality that $\|y_i\|_2 + z_i^T y_i \geq 0$, so the minimum is reached when $y_i = 0$.)

We can minimize over x by setting the gradient with respect to x equal to zero. This yields

$$x = x_0 + \sum_{i=1}^{N} A_i^T z.$$

Substituting in the Lagrangian gives the dual function

$$g(z_1, \ldots, z_N) = \begin{cases} -\sum_{i=1}^N (A_i x_0 + b_i)^T z_i - \frac{1}{2}\|\sum_{i=1}^N A_i^T z_i\|_2^2 & \|z_i\|_2 \leq 1, \ \ i = 1, \ldots, N \\ -\infty & \text{otherwise.} \end{cases}$$

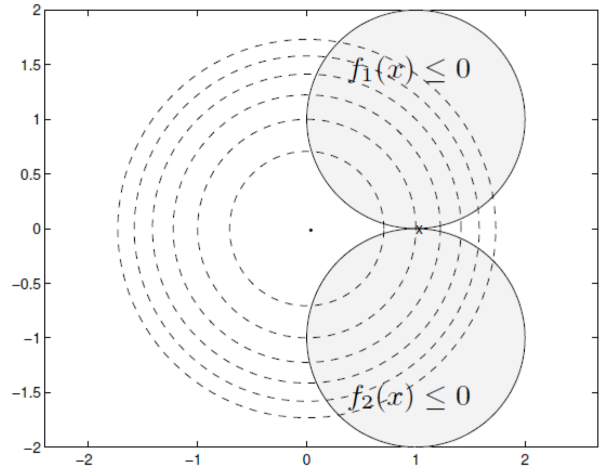
The dual problem is

$$\begin{array}{ll} \text{maximize} & -\sum_{i=1}^N (A_i x_0 + b_i)^T z_i - \frac{1}{2}\|\sum_{i=1}^N A_i^T z_i\|^2 \\ \text{subject to} & \|z_i\|_2 \leq 1, \ \ i = 1, \ldots, N. \end{array}$$

5.26 [H]

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 KKT conditions are

$$
\begin{gathered}
(x_1 - 1)^2 + (x_2 - 1)^2 \leq 1, \quad (x_1 - 1)^2 + (x_2 + 1)^2 \leq 1, \\
\lambda_1 \geq 0, \qquad \lambda_2 \geq 0 \\
2x_1 + 2\lambda_1(x_1 - 1) + 2\lambda_2(x_1 - 1) = 0 \\
2x_2 + 2\lambda_1(x_2 - 1) + 2\lambda_2(x_2 + 1) = 0 \\
\lambda_1((x_1 - 1)^2 + (x_2 - 1)^2 - 1) = \lambda_2((x_1 - 1)^2 + (x_2 + 1)^2 - 1) = 0.
\end{gathered}
$$

At $x = (1, 0)$, these conditions reduce to

$$
\lambda_1 \geq 0, \qquad \lambda_2 \geq 0, \qquad 2 = 0, \qquad -2\lambda_1 + 2\lambda_2 = 0,
$$

which (clearly, in view of the third equation) have no solution.

(c) The Lagrange dual function is given by

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

where

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

L reaches its minimum for

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

and we find

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

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

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

Since g is symmetric, the optimum (if it exists) occurs with $\lambda_1 = \lambda_2$. The dual function then simplifies to

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

We see that $g(\lambda_1, \lambda_2)$ tends to 1 as $\lambda_1 \to \infty$. We have $d^\star = p^\star = 1$, but the dual optimum is not attained.

5.39 [M]

Solution.

(a) Follows from $\mathbf{tr}(Wxx^T) = x^T W x$ and $(xx^T)_{ii} = x_i^2$.

(b) It gives a lower bound because we minimize the same objective over a larger set. If X is rank one, it is optimal.

(c) We write the problem as a minimization problem

$$\begin{array}{ll} \text{minimize} & \mathbf{1}^T \nu \\ \text{subject to} & W + \mathbf{diag}(\nu) \succeq 0. \end{array}$$

Introducing a Lagrange multiplier $X \in \mathbf{S}^n$ for the matrix inequality, we obtain the Lagrangian

$$\begin{aligned} L(\nu, X) &= \mathbf{1}^T \nu - \mathbf{tr}(X(W + \mathbf{diag}(\nu))) \\ &= \mathbf{1}^T \nu - \mathbf{tr}(XW) - \sum_{i=1}^n \nu_i X_{ii} \\ &= -\mathbf{tr}(XW) + \sum_{i=1}^n \nu_i (1 - X_{ii}). \end{aligned}$$

This is bounded below as a function of ν only if $X_{ii} = 1$ for all i, so we obtain the dual problem

$$\begin{array}{ll} \text{maximize} & -\mathbf{tr}(WX) \\ \text{subject to} & X \succeq 0 \\ & X_{ii} = 1, \quad i = 1, \ldots, n. \end{array}$$

Changing the sign again, and switching from maximization to minimization, yields the problem in part (a).