HW1 solutions

Exercise 1 (*Some sets of probability distributions.*) Let x be a real-valued random variable with $\mathbf{Prob}(x = a_i) = p_i$, $i = 1, \ldots, n$, where $a_1 < a_2 < \cdots < a_n$. Of course $p \in \mathbf{R}^n$ lies in the standard probability simplex $P = \{p \mid \mathbf{1}^T p = 1,\ p \succeq 0\}$. Which of the following conditions are convex in p? (That is, for which of the following conditions is the set of $p \in P$ that satisfy the condition convex?)

1. $\alpha \leq \mathbf{E} f(x) \leq \beta$, where $\mathbf{E} f(x)$ is the expected value of $f(x)$, *i.e.*, $\mathbf{E} f(x) = \sum_{i=1}^n p_i f(a_i)$. (The function $f : \mathbf{R} \to \mathbf{R}$ is given.)
2. $\mathbf{Prob}(x > \alpha) \leq \beta$.
3. $\mathbf{E}|x^3| \leq \alpha \mathbf{E}|x|$.
4. $\mathbf{E}x^2 \leq \alpha$.
5. $\mathbf{E}x^2 \geq \alpha$.
6. $\mathbf{var}(x) \leq \alpha$, where $\mathbf{var}(x) = \mathbf{E}(x - \mathbf{E}x)^2$ is the variance of x.
7. $\mathbf{var}(x) \geq \alpha$.
8. $\mathbf{quartile}(x) \geq \alpha$, where $\mathbf{quartile}(x) = \inf\{\beta \mid \mathbf{Prob}(x \leq \beta) \geq 0.25\}$.
9. $\mathbf{quartile}(x) \leq \alpha$.

Solution 1 *We first note that the constraints $p_i \geq 0$, $i = 1, \ldots, n$, define halfspaces, and $\sum_{i=1}^n p_i = 1$ defines a hyperplane, so P is a polyhedron.*

The first five constraints are, in fact, linear inequalities in the probabilities p_i.

1. $\mathbf{E} f(x) = \sum_{i=1}^n p_i f(a_i)$, *so the constraint is equivalent to two linear inequalities*

$$\alpha \leq \sum_{i=1}^n p_i f(a_i) \leq \beta.$$

2. $\mathbf{Prob}(x \geq \alpha) = \sum_{i:\, a_i \geq \alpha} p_i$, *so the constraint is equivalent to a linear inequality*

$$\sum_{i:\, a_i \geq \alpha} p_i \leq \beta.$$

3. The constraint is equivalent to a linear inequality

$$\sum_{i=1}^n p_i (|a_i^3| - \alpha |a_i|) \leq 0.$$

4. *The constraint is equivalent to a linear inequality*

$$\sum_{i=1}^{n} p_i a_i^2 \leq \alpha.$$

5. *The constraint is equivalent to a linear inequality*

$$\sum_{i=1}^{n} p_i a_i^2 \geq \alpha.$$

The first five constraints therefore define convex sets.

6. *The constraint*

$$\mathbf{var}(x) = \mathbf{E}x^2 - (\mathbf{E}x)^2 = \sum_{i=1}^{n} p_i a_i^2 - (\sum_{i=1}^{n} p_i a_i)^2 \leq \alpha$$

is not convex in general. As a counterexample, we can take $n = 2$, $a_1 = 0$, $a_2 = 1$, *and* $\alpha = 1/5$. $p = (1, 0)$ *and* $p = (0, 1)$ *are two points that satisfy* $\mathbf{var}(x) \leq \alpha$, *but the convex combination* $p = (1/2, 1/2)$ *does not.*

7. *This constraint is equivalent to*

$$\sum_{i=1}^{n} a_i^2 p_i + (\sum_{i=1}^{n} a_i p_i)^2 = b^T p + p^T A p \leq \alpha$$

where $b_i = a_i^2$ *and* $A = aa^T$. *This defines a convex set, since the matrix* aa^T *is positive semidefinite.*

Let us denote $\mathbf{quartile}(x) = f(p)$ *to emphasize it is a function of* p.

8. *The constraint* $f(p) \geq \alpha$ *is equivalent to*

$$\mathbf{Prob}(x \leq \beta) < 0.25 \text{ for all } \beta < \alpha.$$

If $\alpha \leq a_1$, *this is always true. Otherwise, define* $k = \max\{i \mid a_i < \alpha\}$. *This is a fixed integer, independent of* p. *The constraint* $f(p) \geq \alpha$ *holds if and only if*

$$\mathbf{Prob}(x \leq a_k) = \sum_{i=1}^{k} p_i < 0.25.$$

This is a strict linear inequality in p, *which defines an open halfspace.*

9. *The constraint* $f(p) \leq \alpha$ *is equivalent to*

$$\mathbf{Prob}(x \leq \beta) \geq 0.25 \text{ for all } \beta \geq \alpha.$$

This can be expressed as a linear inequality

$$\sum_{i=k+1}^{n} p_i \geq 0.25.$$

(If $\alpha \leq a_1$, *we define* $k = 0$.*)*

Exercise 2 (*Euclidean distance matrices.*) Let $x_1, \ldots, x_n \in \mathbf{R}^k$. The matrix $D \in \mathbf{S}^n$ defined by $D_{ij} = \|x_i - x_j\|_2^2$ is called a *Euclidean distance matrix*. It satisfies some obvious properties such as $D_{ij} = D_{ji}$, $D_{ii} = 0$, $D_{ij} \geq 0$, and (from the triangle inequality) $D_{ik}^{1/2} \leq D_{ij}^{1/2} + D_{jk}^{1/2}$. We now pose the question: When is a matrix $D \in \mathbf{S}^n$ a Euclidean distance matrix (for some points in $\mathbf{R}^k$, for some k)? A famous result answers this question: $D \in \mathbf{S}^n$ is a Euclidean distance matrix if and only if $D_{ii} = 0$ and $x^T D x \leq 0$ for all x with $\mathbf{1}^T x = 0$.

Show that the set of Euclidean distance matrices is a convex cone. Find the dual cone.

Solution 2 *The set of Euclidean distance matrices in $\mathbf{S}^n$ is a closed convex cone because it is the intersection of (infinitely many) halfspaces defined by the following homogeneous inequalities:*

$$e_i^T D e_i \leq 0, \qquad e_i^T D e_i \geq 0, \qquad x^T D x = \sum_{j,k} x_j x_k D_{jk} \leq 0,$$

for all $i = 1, \ldots, n$, and all x with $\mathbf{1}^T x = 1$.

It follows that dual cone is given by

$$K^* = \mathbf{Co}(\{-xx^T \mid \mathbf{1}^T x = 1\} \bigcup \{e_1 e_1^T, -e_1 e_1^T, \ldots, e_n e_n^T, -e_n e_n^T\}).$$

This can be made more explicit as follows. Define $V \in \mathbf{R}^{n \times (n-1)}$ as

$$V_{ij} = \begin{cases} 1 - 1/n & i = j \\ -1/n & i \neq j. \end{cases}$$

The columns of V form a basis for the set of vectors orthogonal to $\mathbf{1}$, i.e., a vector x satisfies $\mathbf{1}^T x = 0$ if and only if $x = Vy$ for some y. The dual cone is

$$K^* = \{VWV^T + \mathbf{diag}(u) \mid W \preceq 0, u \in \mathbf{R}^n\}.$$

Exercise 3 (*Composition rules.*) Show that the following functions are convex.

1. $f(x) = -\log(-\log(\sum_{i=1}^m e^{a_i^T x + b_i}))$ 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.
2. $f(x, u, v) = -\sqrt{uv - x^T x}$ 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 3

1. *$g(x) = \log(\sum_{i=1}^m e^{a_i^T x + b_i})$ 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.*
2. *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.*

Exercise 4 (*Problems involving ℓ_1- and ℓ_∞-norms.*) Formulate the following problems as LPs. Explain in detail the relation between the optimal solution of each problem and the solution of its equivalent LP.

1. Minimize $\|Ax - b\|_\infty$ (ℓ_∞-norm approximation).

2. Minimize $\|Ax - b\|_1$ (ℓ_1-norm approximation).

3. Minimize $\|Ax - b\|_1$ subject to $\|x\|_\infty \leq 1$.

4. Minimize $\|x\|_1$ subject to $\|Ax - b\|_\infty \leq 1$.

5. Minimize $\|Ax - b\|_1 + \|x\|_\infty$.

In each problem, $A \in \mathbf{R}^{m \times n}$ and $b \in \mathbf{R}^m$ are given. (See §**??** for more problems involving approximation and constrained approximation.)

Solution 4 ***Solution.***

1. *Equivalent to the LP*

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

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

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

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

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

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.

2. *Equivalent to the LP*

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

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

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

for each k, i.e., $s_k \geq |a_k^T x - b_k|$. 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.

3. *Equivalent to the LP*

$$\begin{array}{ll} \textit{minimize} & \mathbf{1}^T y \\ \textit{subject to} & -y \preceq Ax - b \preceq y \\ & -\mathbf{1} \leq x \leq \mathbf{1}, \end{array}$$

   with variables $x \in \mathbf{R}^n$ *and* $y \in \mathbf{R}^m$.

4. *Equivalent to the LP*

$$\begin{array}{ll} \textit{minimize} & \mathbf{1}^T y \\ \textit{subject to} & -y \leq x \leq y \\ & -\mathbf{1} \leq Ax - b \leq \mathbf{1} \end{array}$$

   with variables x *and* y.

   Another good solution is to write x *as the difference of two nonnegative vectors* $x = x^+ - x^-$, *and to express the problem as*

$$\begin{array}{ll} \textit{minimize} & \mathbf{1}^T x^+ + \mathbf{1}^T x^- \\ \textit{subject to} & -\mathbf{1} \preceq Ax^+ - Ax^- - b \preceq \mathbf{1} \\ & x^+ \succeq 0, \quad x^- \succeq 0, \end{array}$$

   with variables $x^+ \in \mathbf{R}^n$ *and* $x^- \in \mathbf{R}^n$.

5. *Equivalent to*

$$\begin{array}{ll} \textit{minimize} & \mathbf{1}^T y + t \\ \textit{subject to} & -y \preceq Ax - b \preceq y \\ & -t\mathbf{1} \preceq x \preceq t\mathbf{1}, \end{array}$$

   with variables x, y, *and* t.

Exercise 5 (*Linear separation of two sets of ellipsoids.*) Suppose we are given $K + L$ ellipsoids

$$\mathcal{E}_i = \{P_i u + q_i \mid \|u\|_2 \leq 1\}, \quad i = 1, \ldots, K + L,$$

where $P_i \in \mathbf{S}^n$. We are interested in finding a hyperplane that strictly separates $\mathcal{E}_1, \ldots, \mathcal{E}_K$ from $\mathcal{E}_{K+1}, \ldots, \mathcal{E}_{K+L}$, *i.e.*, we want to compute $a \in \mathbf{R}^n$, $b \in \mathbf{R}$ such that

$$a^T x + b > 0 \text{ for } x \in \mathcal{E}_1 \cup \cdots \cup \mathcal{E}_K, \qquad a^T x + b < 0 \text{ for } x \in \mathcal{E}_{K+1} \cup \cdots \cup \mathcal{E}_{K+L},$$

or prove that no such hyperplane exists. Express this problem as an SOCP feasibility problem.

Solution 5 ***Solution.*** *We first note that the problem is homogeneous in* a *and* b, *so we can replace the strict inequalities* $a^T x + b > 0$ *and* $a^T x + b < 0$ *with* $a^T x + b \geq 1$ *and* $a^T x + b \leq -1$, *respectively.*

The variables a *and* b *must satisfy*

$$\inf_{\|u\|_2 \leq 1} (a^T P_i u + a^T q_i) \geq 1, \quad 1, \ldots, L$$

and

$$\sup_{\|u\|_2\leq 1} (a^T P_i u + a^T q_i) \leq -1, \quad i = K+1, \ldots, K+L.$$

The lefthand sides can be expressed as

$$\inf_{\|u\|_2\leq 1} (a^T P_i u + a^T q_i) = -\|P_i^T a\|_2 + a^T q_i + b, \qquad \sup_{\|u\|_2\leq 1} (a^T P_i u + a^T q_i) = \|P_i^T a\|_2 + a^T q_i + b.$$

We therefore obtain a set of second-order cone constraints in a, b*:*

$$\begin{array}{ll} -\|P_i^T a\|_2 + a^T q_i + b \geq 1, & i = 1, \ldots, L \\ \|P_i^T a\|_2 + a^T q_i + b \leq -1, & i = K+1, \ldots, K+L. \end{array}$$

Exercise 6 (*Dual of general LP.*) Find the dual function of the LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Gx \preceq h \\ & Ax = b. \end{array}$$

Give the dual problem, and make the implicit equality constraints explicit.

Solution 6 ***Solution.***

1. *The Lagrangian is*

$$\begin{array}{rcl} L(x, \lambda, \nu) & = & c^T x + \lambda^T (Gx - h) + \nu^T (Ax - b) \\ & = & (c^T + \lambda^T G + \nu^T A)x - h\lambda^T - \nu^T b, \end{array}$$

   which is an affine function of x*. It follows that the dual function is given by*

$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu) = \left\{ \begin{array}{ll} -\lambda^T h - \nu^T b & c + G^T \lambda + A^T \nu = 0 \\ -\infty & \textit{otherwise.} \end{array} \right.$$

2. *The dual problem is*

$$\begin{array}{ll} \textit{maximize} & g(\lambda, \nu) \\ \textit{subject to} & \lambda \succeq 0. \end{array}$$

   After making the implicit constraints explicit, we obtain

$$\begin{array}{ll} \textit{maximize} & -\lambda^T h - \nu^T b \\ \textit{subject to} & c + G^T \lambda + A^T \nu = 0 \\ & \lambda \succeq 0. \end{array}$$