Solutions HW #2

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.

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

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

2. The dual problem is

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

After making the implicit constraints explicit, we obtain

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

Piecewise-linear minimization.

We consider the convex piecewise-linear minimization problem

$$\text{minimize} \quad \max_{i=1,\ldots,m}(a_i^T x + b_i) \tag{1}$$

with variable $x \in \mathbf{R}^n$.

1. Derive a dual problem, based on the Lagrange dual of the equivalent problem

$$\begin{array}{ll} \text{minimize} & \max_{i=1,\ldots,m} y_i \\ \text{subject to} & a_i^T x + b_i = y_i, \quad i = 1, \ldots, m, \end{array}$$

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

2. Formulate the piecewise-linear minimization problem (1) as an LP, and form the dual of the LP. Relate the LP dual to the dual obtained in part (a).

3. Suppose we approximate the objective function in (1) by the smooth function

$$f_0(x) = \log\left(\sum_{i=1}^m \exp(a_i^T x + b_i)\right),$$

and solve the unconstrained geometric program

$$\text{minimize} \quad \log\left(\sum_{i=1}^m \exp(a_i^T x + b_i)\right). \tag{2}$$

A dual of this problem is given by (5.62). Let $p^\star_{\mathrm{pwl}}$ and $p^\star_{\mathrm{gp}}$ be the optimal values of (1) and (2), respectively. Show that

$$0 \le p^\star_{\mathrm{gp}} - p^\star_{\mathrm{pwl}} \le \log m.$$

4. Derive similar bounds for the difference between $p^\star_{\mathrm{pwl}}$ and the optimal value of

$$\text{minimize} \quad (1/\gamma)\log\left(\sum_{i=1}^m \exp(\gamma(a_i^T x + b_i))\right),$$

where $\gamma > 0$ is a parameter. What happens as we increase γ?

Solution.

1. The dual function is

$$g(\lambda) = \inf_{x,y}\left(\max_{i=1,\dots,m} y_i + \sum_{i=1}^m \lambda_i(a_i^T x + b_i - y_i)\right).$$

The infimum over x is finite only if $\sum_i \lambda_i a_i = 0$. To minimize over y we note that

$$\inf_y(\max_i y_i - \lambda^T y) = \begin{cases} 0 & \lambda \succeq 0, \quad \mathbf{1}^T\lambda = 1 \\ -\infty & \text{otherwise.} \end{cases}$$

To prove this, we first note that if $\lambda \succeq 0$, $\mathbf{1}^T\lambda = 1$, then

$$\lambda^T y = \sum_j \lambda_j y_j \le \sum_j \lambda_j \max_i y_i = \max_i y_i,$$

with equality if $y = 0$, so in that case

$$\inf_y(\max_i y_i - \lambda^T y) = 0.$$

If $\lambda \not\succeq 0$, say $\lambda_j < 0$, then choosing $y_i = 0$, $i \ne j$, and $y_j = -t$, with $t \ge 0$, and letting t go to infinity, gives

$$\max_i y_i - \lambda^T y = 0 + t\lambda_k \to -\infty.$$

Finally, if $\mathbf{1}^T\lambda \ne 1$, choosing $y = t\mathbf{1}$, gives

$$\max_i y_i - \lambda^T y = t(1 - \mathbf{1}^T\lambda) \to -\infty,$$

if $t \to \infty$ and $1 < \mathbf{1}^T\lambda$, or if $t \to -\infty$ and $1 > \mathbf{1}^T\lambda$.

Summing up, we have

$$g(\lambda) = \begin{cases} b^T\lambda & \sum_i \lambda_i a_i = 0, \quad \lambda \succeq 0, \quad \mathbf{1}^T\lambda = 1 \\ -\infty & \text{otherwise.} \end{cases}$$

The resulting dual problem is

$$\begin{array}{ll} \text{maximize} & b^T\lambda \\ \text{subject to} & A^T\lambda = 0 \\ & \mathbf{1}^T\lambda = 1 \\ & \lambda \succeq 0. \end{array}$$

2. The problem is equivalent to the LP

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

The dual problem is

$$\begin{array}{ll} \text{maximize} & b^T z \\ \text{subject to} & A^T z = 0, \quad \mathbf{1}^T z = 1, \quad z \succeq 0, \end{array}$$

which is identical to the dual derived in (a).

3. First both primal problems PWL and GP are convex so strong duality holds.

Suppose $z^\star$ is dual optimal for the dual GP (5.62) (with the convention $0 \log 0 = 0$)

$$\begin{array}{ll} \text{maximize} & b^T z - \sum_{i=1}^m z_i \log z_i \\ \text{subject to} & \mathbf{1}^T z = 1 \\ & A^T z = 0 \\ & z \succeq 0 \end{array}$$

Then $z^\star$ is also feasible for the dual of the piecewise-linear formulation, with objective value

$$b^T z = p^\star_{\mathrm{gp}} + \sum_{i=1}^m z_i^\star \log z_i^\star.$$

This provides a lower bound on $p^\star_{\mathrm{pwl}}$:

$$p^\star_{\mathrm{pwl}} \geq p^\star_{\mathrm{gp}} + \sum_{i=1}^m z_i^\star \log z_i^\star \geq p^\star_{\mathrm{gp}} - \log m.$$

The bound follows from concavity of log using Jensens's inequality

$$\sum_{i=1}^m z_i \log(1/z_i) \leq \log \sum_{i=1}^m 1 = \log m.$$

On the other hand, we also have

$$\max_i (a_i^T x + b_i) \leq \log \sum_i \exp(a_i^T x + b_i)$$

for all x, and therefore $p^\star_{\mathrm{pwl}} \leq p^\star_{\mathrm{gp}}$.

In conclusion,

$$p^\star_{\mathrm{gp}} - \log m \leq p^\star_{\mathrm{pwl}} \leq p^\star_{\mathrm{gp}}.$$

4. We first reformulate the problem as

$$\begin{array}{ll} \text{minimize} & (1/\gamma)\log\sum_{i=1}^m \exp(\gamma y_i) \\ \text{subject to} & Ax+b=y. \end{array}$$

The Lagrangian is

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

L is bounded below as a function of x only if $A^T z = 0$. To find the optimum over y, we compute the conjugate function for $z \in \mathbf{R}^m$,

$$f^*(z) = \sup_y \gamma z^T y - \log\sum_{i=1}^m \exp(\gamma y_i) = \sup_y\ z^T y - \log\sum_{i=1}^m \exp(y_i).$$

If there exists $z_i < 0$, then, taking $y = te_i$, we get

$$z^T y - \log\sum_{i=1}^m \exp(y_i) = tz_i - \log((m-1)+e^t),$$

which goes to infinity for $t \to -\infty$.

If $z \succeq 0$ and $\mathbf{1}^T z \neq 1$, then, taking $y = t\mathbf{1}$, we get

$$z^T y - \log\sum_{i=1}^m \exp(y_i) = t\mathbf{1}^T z - \log(m) - t,$$

which goes to infinity for $t \to +\infty$ or $-\infty$ depending on the sign of $\mathbf{1}^T z - 1$.

If $z \succeq 0$ and $\mathbf{1}^T z = \sum_{z_i \neq 0} z_i = 1$, we have using concavity of log,

$$z^T y - \sum_{i=1}^m z_i \log z_i = \sum_{z_i\neq 0} z_i(y_i - \log z_i) \leq \log\sum_{z_i\neq 0} z_i \exp(y_i - \log(z_i)) \leq \log\sum_{i=1}^m \exp(y_i),$$

with equality by taking $y_i = \log z_i$ if $z_i > 0$ and $y_i \to -\infty$ for $z_i = 0$. Therefore we have

$$f^*(z) = \sum_{i=1}^m z_i \log z_i.$$

The Lagrange dual function is then given for $z \succeq 0,\ \mathbf{1}^T z = 1,\ A^T z = 0$ by

$$g(z) = b^T z - \frac{1}{\gamma}\sum_{i=1}^m z_i \log z_i,$$

and the dual problem is

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

Let $p^\star_{\rm gp}(\gamma)$ be the optimal value of the GP. Following the same argument as above, we can conclude that

$$p^\star_{\rm gp}(\gamma) - \frac{1}{\gamma}\log m \le p^\star_{\rm pwl} \le p^\star_{\rm gp}(\gamma).$$

In other words, $p^\star_{\rm gp}(\gamma)$ approaches $p^\star_{\rm pwl}$ as γ increases.

Suboptimality of a simple covering ellipsoid.

Recall the problem of determining the minimum volume ellipsoid, centered at the origin, that contains the points $a_1, \ldots, a_m \in \mathbf{R}^n$ (problem (5.14), page 222):

$$\begin{array}{ll} \text{minimize} & f_0(X) = \log\det(X^{-1}) \\ \text{subject to} & a_i^T X a_i \le 1, \quad i = 1, \ldots, m, \end{array}$$

with $\mathbf{dom}\, f_0 = \mathbf{S}^n_{++}$. We assume that the vectors $a_1, \ldots, a_m$ span $\mathbf{R}^n$ (which implies that the problem is bounded below).

1. Show that the matrix

$$X_{\rm sim} = \left(\sum_{k=1}^m a_k a_k^T\right)^{-1},$$

is feasible. *Hint.* Show that

$$\left[\begin{array}{cc} \sum_{k=1}^m a_k a_k^T & a_i \\ a_i^T & 1 \end{array}\right] \succeq 0,$$

and use Schur complements (§A.5.5) to prove that $a_i^T X a_i \le 1$ for $i = 1, \ldots, m$.

Solution.

$$\left[\begin{array}{cc} \sum_{k=1}^m a_k a_k^T & a_k \\ a_i^T & 1 \end{array}\right] = \left[\begin{array}{cc} \sum_{k\neq i} a_k a_k^T & 0 \\ 0 & 0 \end{array}\right] + \left[\begin{array}{c} a_i \\ 1 \end{array}\right]\left[\begin{array}{c} a_i \\ 1 \end{array}\right]^T$$

is the sum of two positive semidefinite matrices, hence positive semidefinite. The Schur complement of the 1,1 block of this matrix, $\sum_{k=1}^m a_k a_k^T$ which is invertible by hypothesis, is therefore also positive semidefinite:

$$1 - a_i^T \left(\sum_{k=1}^m a_k a_k^T\right)^{-1} a_i \ge 0,$$

which is the desired conclusion.

2. Now we establish a bound on how suboptimal the feasible point $X_{\rm sim}$ is, via the dual problem,

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

with the implicit constraint $\sum_{i=1}^m \lambda_i a_i a_i^T \succ 0$. (This dual is derived on page 222.)

To derive a bound, we restrict our attention to dual variables of the form $\lambda = t\mathbf{1}$, where $t > 0$. Find (analytically) the optimal value of t, and evaluate the dual objective at this λ. Use this to prove that the volume of the ellipsoid $\{u \mid u^T X_{\rm sim} u \le 1\}$ is no more than a factor $(m/n)^{n/2}$ more than the volume of the minimum volume ellipsoid.

Solution. The dual function evaluated at $\lambda = t\mathbf{1}$ is

$$g(\lambda) = \log\det\left(\sum_{i=1}^m a_i a_i^T\right) + n\log t - mt + n.$$

Now we'll maximize over $t > 0$ to get the best lower bound. Setting the derivative with respect to t equal to zero yields the optimal value $t = n/m$. Using this λ we get the dual objective value

$$g(\lambda) = \log\det\left(\sum_{i=1}^m a_i a_i^T\right) + n\log(n/m).$$

The primal objective value for $X = X_{\rm sim}$ is given by

$$-\log\det\left(\sum_{i=1}^m a_i a_i^T\right)^{-1},$$

so the duality gap associated with $X_{\rm sim}$ and λ is $n\log(m/n)$. (Recall that $m \geq n$, by our assumption that $a_1, \ldots, a_m$ span $\mathbf{R}^n$.) It follows that, in terms of the objective function, $X_{\rm sim}$ is no more than $n\log(m/n)$ suboptimal. The volume V of the ellipsoid $\mathcal{E}$ associated with the matrix X is given by $V = \exp(-O/2)$, where O is the associated objective function, $O = -\log\det X$. The bound follows.

Dual problem.

Derive a dual problem for

$$\text{minimize} \quad \textstyle\sum_{i=1}^N \|A_i x + b_i\|_2 + (1/2)\|x - x_0\|_2^2.$$

The problem data are $A_i \in \mathbf{R}^{m_i \times n}$, $b_i \in \mathbf{R}^{m_i}$, and $x_0 \in \mathbf{R}^n$. First introduce new variables $y_i \in \mathbf{R}^{m_i}$ and equality constraints $y_i = A_i x + b_i$.

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

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

The dual problem is

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

Lagrangian relaxation of Boolean LP.

A *Boolean linear program* is an optimization problem of the form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \\ & x_i \in \{0, 1\}, \quad i = 1, \ldots, n, \end{array}$$

and is, in general, very difficult to solve. In exercise (4.15) we studied the LP relaxation of this problem,

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \\ & 0 \leq x_i \leq 1, \quad i = 1, \ldots, n, \end{array} \tag{3}$$

which is far easier to solve, and gives a lower bound on the optimal value of the Boolean LP. In this problem we derive another lower bound for the Boolean LP, and work out the relation between the two lower bounds.

1. *Lagrangian relaxation.* The Boolean LP can be reformulated as the problem

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax \preceq b \\ & x_i(1 - x_i) = 0, \quad i = 1, \ldots, n, \end{array}$$

   which has quadratic equality constraints. Find the Lagrange dual of this problem. The optimal value of the dual problem (which is convex) gives a lower bound on the optimal value of the Boolean LP. This method of finding a lower bound on the optimal value is called *Lagrangian relaxation.*

2. Show that the lower bound obtained via Lagrangian relaxation, and via the LP relaxation (3), are the same. *Hint.* Derive the dual of the LP relaxation (3).

Solution.

1. The Lagrangian is

$$\begin{array}{rcl} L(x, \mu, \nu) & = & c^T x + \mu^T (Ax - b) - \nu^T x + x^T \mathbf{diag}(\nu) x \\ & = & x^T \mathbf{diag}(\nu) x + (c + A^T \mu - \nu)^T x - b^T \mu. \end{array}$$

   Minimizing over x gives the dual function

$$g(\mu, \nu) = \begin{cases} -b^T \mu - (1/4) \sum_{i=1}^n (c_i + a_i^T \mu - \nu_i)^2 / \nu_i & \nu \succeq 0 \\ -\infty & \text{otherwise} \end{cases}$$

   where a_i is the ith column of A, and we adopt the convention that $a^2/0 = \infty$ if $a \neq 0$, and $a^2/0 = 0$ if $a = 0$.

The resulting dual problem is

$$\begin{array}{ll} \text{maximize} & -b^T\mu - (1/4)\sum_{i=1}^n (c_i + a_i^T\mu - \nu_i)^2/\nu_i \\ \text{subject to} & \nu \succeq 0. \end{array}$$

In order to simplify this dual, we optimize analytically over ν, by noting that

$$\begin{array}{rcl} \sup_{\nu_i \geq 0} \left(-\frac{(c_i + a_i^T\mu - \nu_i)^2}{\nu_i} \right) & = & \left\{ \begin{array}{ll} 4(c_i + a_i^T\mu) & c_i + a_i^T\mu \leq 0 \\ 0 & c_i + a_i^T\mu \geq 0 \end{array} \right. \\ & = & 4\min\{0, (c_i + a_i^T\mu)\} \end{array}$$

This allows us to eliminate ν from the dual problem, and simplify it as

$$\begin{array}{ll} \text{maximize} & -b^T\mu + \sum_{i=1}^n \min\{0, c_i + a_i^T\mu\} \\ \text{subject to} & \mu \succeq 0. \end{array}$$

2. We follow the hint. The Lagrangian and dual function of the LP relaxation are

$$\begin{array}{rcl} L(x,u,v,w) & = & c^Tx + u^T(Ax-b) - v^Tx + w^T(x-\mathbf{1}) \\ & = & (c + A^Tu - v + w)^Tx - b^Tu - \mathbf{1}^Tw \\ g(u,v,w) & = & \left\{ \begin{array}{ll} -b^Tu - \mathbf{1}^Tw & A^Tu - v + w + c = 0 \\ -\infty & \text{otherwise.} \end{array} \right. \end{array}$$

The dual problem is

$$\begin{array}{ll} \text{maximize} & -b^Tu - \mathbf{1}^Tw \\ \text{subject to} & A^Tu - v + w + c = 0 \\ & u \succeq 0, v \succeq 0, w \succeq 0, \end{array}$$

which is equivalent to the Lagrange relaxation problem derived above. We conclude that the two relaxations give the same value.