[M2, MVA]
Convex Optimization, Algorithms and Applications

Maha ELBAYAD
maha.elbayad@student.ecp.fr

Homework 2

October 25, 2015

Exercise 5.5 - Dual of general LP.

We consider the LP:

$$\begin{aligned} \underset{x}{\text{minimize}} \quad & c^T x \\ \text{subject to} \quad & Gx \preceq h \\ & Ax = b \end{aligned} \tag{1}$$

The Lagrangien of (1) is:

$$L(x, \lambda, \nu) = c^T x + \lambda^T (Gx - h) + \nu^T (Ax - b) = (c^T + \lambda^T G + \nu^T A)x - \lambda^T h - \nu^T b.$$

The corresponding Lagrange dual function is:

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

Hence the dual problem:

$$\begin{aligned} \text{maximize} \quad & g(\lambda, \nu) \\ \text{subject to} \quad & \lambda \succeq \mathbf{0} \end{aligned}$$

With explicit equality contraints:

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

Exercise 5.7 - Piecewise-linear minimization

We consider the convex piecewise-linear minimization problem:

$$Minimize \qquad \max_{i=1,...,m} (a_i^T x + b_i) \tag{2}$$

with $x \in \mathbb{R}^n$.

(a) Let us consider the equivalent problem:

$$\begin{aligned} &\text{minimize} \qquad \max_{i=1,...,m} y_i \\ &\text{subject to} \quad y_i = a_i^T x + b_i, \ i = 1, ..., m. \end{aligned} \tag{3}$$

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

The dual function of (3) is as follows:

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

We minimize for x then for y:

$$\inf_x \left(\max_{i=1,...,m} y_i + \sum_{i=1}^m \lambda_i (a_i^T x + b_i - y_i \right) = \begin{cases} \max_i y_i - \lambda^T y + \lambda^T b & if \ \sum\limits_{i=1}^m \lambda_i a_i^T = 0 \\ -\infty \quad otherwise \end{cases}$$

Then:

$$g(\lambda) = \inf_y (\max_i y_i - \lambda^T y + \lambda^T b) \qquad if A^T \lambda = 0.$$

If λ has a nonpositive component say $\lambda_k < 0$ then for $y = -\alpha e_k$ whith $\alpha \geq 0$:

$$\max_i y_i - \lambda^T y + \lambda^T b = \alpha \lambda_k + \lambda^T b \xrightarrow[\alpha \to +\infty]{} -\infty$$

If $\mathbf{1}^T y \neq 1$ for $y = \alpha \mathbf{1}$ ($\alpha \in \mathbb{R}$):

$$\max_i y_i - \lambda^T y + \lambda^T b = \alpha(1 - \mathbf{1}^T \lambda) + \lambda^T b \xrightarrow[\alpha \to \pm\infty]{} -\infty$$

depending on the sign of $\mathbf{1}^T \lambda - 1$.

If $\mathbf{1}^T \lambda = 1 \ \wedge \ \lambda \succeq \mathbf{0}$ then:

$$\lambda^T y \leq \sum_j \lambda_j \max_i y_i = \max_i y_i$$

Hence $\max_i y_i - \lambda^T y + \lambda^T b \geq \lambda^T b$ with equality reached for $y = \mathbf{0}$.

Therefore:

$$g(\lambda) = \begin{cases} \lambda^T b & if \ (A^T \lambda = 0) \wedge (\lambda \succeq \mathbf{0}) \wedge (\mathbf{1}^T \lambda = 1) \\ -\infty & otherwise. \end{cases}$$

The dual problem with explicit constraints is:

$$\begin{aligned} \text{maximize} \quad & \lambda^T b \\ \text{subject to} \quad & A^T\lambda = \mathbf{0} \\ & \mathbf{1}^T\lambda = 1 \\ & \lambda \succeq \mathbf{0} \end{aligned} \tag{4}$$

(b) The problem in (2) is equivalent to the LP:

$$\begin{aligned} \text{minimize} \quad & t \\ \text{subject to} \quad & Ax + b \preceq t\mathbf{1} \end{aligned} \tag{5}$$

The dual function is:

$$g(\lambda) = \inf_{t,x} t(1 - \lambda^T\mathbf{1}) + \lambda^T Ax + \lambda^T b = \begin{cases} \lambda^T b & if\ (\lambda^T\mathbf{1} = 1)\ \wedge\ (\lambda^T A = \mathbf{0})\ \wedge\ (\lambda \succeq \mathbf{0}) \\ -\infty & otherwise \end{cases}$$

Hence, the problem (5) has for dual:

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

Which is the same as (4).

(c) Suppose we approximate the objective function (2) 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:

$$\text{minimize } f_0(x) \tag{6}$$

The dual of this problem is:

$$\begin{aligned} \text{maximize} \quad & b^T\nu - \sum_{i=1}^m \nu_i \log \nu_i \\ \text{subject to} \quad & \mathbf{1}^T\nu = 1 \\ & A^T\nu = \mathbf{0} \\ & \nu \succeq \mathbf{0} \end{aligned} \tag{7}$$

Let p_{pwl}^* and p_{gp}^* be the optimal values of (2) and (6) respectively.
We want to show that:

$$0 \leq p_{gp}^* - p_{pwl}^* \leq \log m$$

For the LHS, we have for all $x \in \mathbb{R}^n$:

$$max_i(a_i^T x + b_i) \leq \log \sum_{i=1}^{m} \exp(a_i^T x + b_i)$$

hence:

$$p_{pwl}^* \leq p_{gp}^*$$

For the RHS, let ν^* be the optimal of the dual (7), ν^* is feasible for (4) with:

$$b^T \nu^* - \sum_{i=1}^{m} \nu_i^* \log \nu_i^* = p_{gp}^*$$

Thus:

$$\nu^{*T} b \leq p_{pwl}^*$$

We can easily show that:

$$\inf_{s.t\ \mathbf{1}^T \nu = 1} \sum_{i=1}^{m} \nu_i \log \nu_i = -\log m$$

Hence:

$$p_{pwl}^* \geq p_{gp}^* + \sum_{i=1}^{m} \nu_i^* \log \nu_i^* \geq p_{gp}^* - \log m$$

Therefore:

$$0 \leq p_{gp}^* - p_{pwl}^* \leq \log m$$

(d) Let us consider the problem:

$$\text{minimize} \quad \frac{1}{\gamma} \log \left(\sum_{i=1}^{m} \exp(\gamma(a_i^T x + b_i)) \right) \tag{8}$$

Where $\gamma > 0$ is a parameter.
The problem in (8) is equivalent to:

$$\begin{aligned} &\text{minimize} \quad \frac{1}{\gamma} \log \left(\sum_{i=1}^{m} \exp(\gamma y_i) \right) \\ &\text{subject to} \quad Ax + b = y \end{aligned}$$

which has the Lagrangian:

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

To compute the dual function, we minimize with respect to x:

$$\inf_x L(x, y, \lambda) = \frac{1}{\gamma} \log \left(\sum_{i=1}^{m} \exp(\gamma y_i) \right) + \lambda^T b - \lambda^T y, if \lambda^T A = \mathbf{0} \text{ and is unbounded otherwise.}$$

For y we solve the following:

$$\frac{\partial}{\partial y_i} (\frac{1}{\gamma} \log \left(\sum_{i=1}^{m} \exp(\gamma y_i) \right) - \lambda^T y) = \frac{\exp(\gamma y_i)}{\sum_{j=1}^{m} \exp(\gamma y_j)} - \lambda_i = 0$$

Which means $\lambda \succeq \mathbf{0}$ and $\mathbf{1}^T \lambda = 1$ with:

$$\lambda^T y = \frac{1}{\gamma} \sum_i \lambda_i \log \lambda_i + \frac{1}{\gamma} \log \left(\sum_{i=1}^{m} \exp(\gamma y_i) \right)$$

Therefore:

$$g(\lambda) = \inf_{x,y} L(x, y, \lambda) = \begin{cases} \lambda^T b - \frac{1}{\gamma} \sum_i \lambda_i \log \lambda_i & if \ (A^T \lambda = 0) \wedge (\mathbf{1}^T \lambda = 1) \wedge (\lambda \succeq \mathbf{0}) \\ -\infty & otherwise \end{cases}$$

And the explicit dual problem:

$$\begin{aligned} \text{maximize} \quad & b^T \lambda - \frac{1}{\gamma} \sum_i \lambda_i \log \lambda_i \\ \text{subject to} \quad & \mathbf{1}^T \lambda = 1 \\ & A^T \lambda = \mathbf{0} \\ & \lambda \succeq \mathbf{0} \end{aligned} \tag{9}$$

Let $p^*(\gamma)$ denote the optimal value of (8) and show that:

$$p^*(\gamma) - \frac{1}{\gamma} \log m \leq p^*_{pwl} \leq p^*(\gamma)$$

For the RHS:

$$max_i(a_i^T x + b_i) \leq \frac{1}{\gamma} \log \sum_{i=1}^{m} \exp(\gamma (a_i^T x + b_i)) \ (\forall x)$$

Thus, $p^*_{pwl} \leq p^*(\gamma)$.
Following the same reasoning in (c) for the LHS and considering λ^* the optimal value of (9) which is feasible for (4):

$$b^T\lambda^* - \frac{1}{\gamma}\sum_{i=1}^{m} \lambda_i^* \log \lambda_i^* = p^*(\gamma)$$

and:

$$\lambda^{*T} b \leq p^*_{pwl}$$

And given that

$$\inf_{s.t\ \mathbf{1}^T\lambda=1} \frac{1}{\gamma}\sum_{i=1}^{m} \lambda_i \log \lambda_i = -\frac{\log m}{\gamma}$$

We find that:

$$p^*(\gamma) - \frac{1}{\gamma}\log m \leq p^*_{pwl} \leq p^*(\gamma)$$

Hence: $p^*(\gamma) \xrightarrow[\gamma\to+\infty]{} p^*_{pwl}$

Exercise 5.9 - Suboptimality of a simple covering ellipsoid

Let us consider the problem of determining the mnimum volume ellipsoid, centered at the origin, that contains the points $a_1, ..., a_m \in \mathbb{R}^n$:

$$\begin{array}{ll} \text{minimize} & f_0(X) = \log\det(X^{-1}) \\ \text{subject to} & a_i^T X a_i \leq 1,\ i = 1, .., m \end{array} \tag{10}$$

With $\mathbf{dom} f_0 = \mathbf{S}^n_{++}$. We assume that $a_1, .., a_m$ span $\mathbb{R}^n$.
(a) Let $X_{sim} = \left(\sum\limits_{k=1}^{m} a_k a_k^T\right)^{-1}$.
To show that X_{sim} is feasible we will show that:

$$\begin{bmatrix} \sum\limits_{k=1}^{m} a_k a_k^T & a_i \\ a_i^T & 1 \end{bmatrix} \succeq \mathbf{0}$$

We have:

$$\begin{bmatrix} \sum\limits_{k=1}^{m} a_k a_k^T & a_i \\ a_i^T & 1 \end{bmatrix} = \begin{bmatrix} \sum\limits_{k\neq i}^{m} a_k a_k^T & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} a_i \\ 1 \end{bmatrix} \begin{bmatrix} a_i^T & 1 \end{bmatrix} \succeq \mathbf{0}$$

As sum of two matrices from $\mathbf{S}^n_+$.
The Schur complement of $\sum\limits_{k=1}^{m} a_k a_k^T$ is positive, hence:

$$S = 1 - a_i^T X_{sim} a_i \geq 0$$

$$a_i^T X_{sim} a_i \leq 1 \quad Q.E.D$$

(b) The dual problem of (10) is:

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

For $\lambda = t\mathbf{1}$, $(t \geq 0)$ the dual function is as follows:

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

we maximize over $t > 0$:

$$\frac{\partial g(t\mathbf{1})}{\partial t} = \frac{n}{t} - m = 0$$

Thus $t^* = \dfrac{n}{m}$.
with:

$$g(t^*\mathbf{1}) = \log\det(\sum_{i=1}^m a_i a_i^T) + n\log(\frac{n}{m})$$

For$X = X_{sim}$ the objective function is:

$$f_0(X_{sim}) = \log\det(\sum_{i=1}^m a_i a_i^T)$$

Hence the duality gap between X_{sim} and $\lambda = t^*\mathbf{1}$ is equal to

$$gap = n\log(n/m) \leq 0, \text{ } (n \leq m \text{ since the m vectors span } \mathbb{R}^n)$$

This is to say that for X^* the optimal solution:

$$f_0(X_{sim}) - f_0(X^*) \leq n\log(\frac{m}{n})$$

We know that for an ellipsoid $\mathcal{E}_X = \{z | \, z^T X z \leq 1\}$, the volume V_X is proportional to $(\exp f_0(X))^{1/2}$:
Thus:

$$\frac{V_{sim}}{V^*} \leq \left(\frac{m}{n}\right)^{n/2}$$

Exercise 5.11 - Dual problem
Consider the problem

$$\text{minimize} \quad \sum_{i=1}^{N} \|A_i x + b_i\|_2 + \frac{1}{2}\|x - x_0\|_2^2 \tag{11}$$

The problem data are $A_i \in \mathbb{R}^{m_i \times n}$, $b_i \in \mathbb{R}^{m_i}$ and $x_0 \in \mathbb{R}^n$.
The problem in (11) is equivalent to:

$$\begin{aligned} &\text{minimize} \quad && \sum_{i=1}^{N} \|y_i\|_2 + \frac{1}{2}\|x - x_0\|_2^2 \\ &\text{subject to} \quad && y_i = A_i x + b_i \ \forall i = 1, .., N \end{aligned}$$

The Lagrangian is:

$$L(x, \lambda_1, ..., \lambda_N) = \sum_{i=1}^{N} \|y_i\|_2 + \frac{1}{2}\|x - x_0\|_2^2 + \sum_{i=1}^{N} \lambda_i^T (y_i - A_i x - b_i)$$

To minimize over y_i, we impose $\|\lambda_i\|_2 \leq 1$.
In fact, if $\|\lambda_i\|_2 > 1$ then for $y_i = -\alpha\lambda_i$, $\alpha > 0$ the component $(\|y_i\|_2 + \lambda_i^T y_i)$ of the Lagrangian is equal to:

$$\alpha\|\lambda_i\|(1 - \|\lambda_i\|) \xrightarrow[\alpha \to +\infty]{} -\infty$$

And if $\|\lambda_i\|_2 \leq 1$ is satisfied, then C.S gives us:

$$\lambda_i^T y_i \geq -\|\lambda_i\|\|y_i\| \Rightarrow \|y_i\|_2 + \lambda_i^T y_i \geq 0$$

With equality if $y_i = \mathbf{0}$, hence:

$$\inf_{y_i} \|y_i\|_2 + \lambda_i^T y_i = \begin{cases} 0, & if \ \|\lambda_i\| \leq 1 \\ -\infty \ otherwise \end{cases}$$

To minimize over x, we set the gradient to 0:

$$\frac{\partial}{\partial x} L(...) = x - x_0 - \sum_{i=1}^{N} A_i^T \lambda_i = \mathbf{0}$$

Thus, the dual function is:

$$g(\lambda_1, ..., \lambda_N) = \begin{cases} -\frac{1}{2}\|\sum_{i=1}^{N} A_i^T \lambda_i\|^2 - \sum_{i=1}^{N} \lambda_i^T (A_i x_0 + b_i), & if \ \|\lambda_i\|_2 \leq 1 \ (i = 1, .., N) \\ -\infty \quad otherwise \end{cases}$$

The explicit dual problem is:

$$\begin{aligned} &\text{maximize} \quad -\frac{1}{2}\|\sum_{i=1}^{N} A_i^T\lambda_i\|^2 - \sum_{i=1}^{N}\lambda_i^T(A_ix_0+b_i) \\ &\text{subject to} \quad \|\lambda_i\| \leq 1 \ \ i=1,..,N \end{aligned}$$

Exercise 5.13 - Lagrangian relaxation of Boolean LP

We consider the Boolean LP:

$$\begin{aligned} &\text{minimize} && c^Tx \\ &\text{subject to} && Ax \preceq b \\ & && x_i \in \{0,1\} \ \ i=1,..,n \end{aligned} \tag{12}$$

(a) *Lagrangian relaxation*:

The problem (12) can be reformulated as:

$$\begin{aligned} &\text{minimize} && c^Tx \\ &\text{subject to} && Ax \preceq b \\ & && x_i(1-x_i) = 0 \ \ i=1,..,n \end{aligned}$$

The Lagrangian is:

$$L(x,\nu,\lambda) = c^Tx + \lambda^T(Ax-b) - \sum_{i=1}^{n}\nu_i x_i(1-x_i) = c^Tx + \lambda^T(Ax-b) - \nu^Tx + x^T Diag(\nu)x$$

Minimizing with respect to x:

If ν has a nonpositive component ν_i then the Lagrangian is unbounded for $x = \alpha e_i$ as $-\nu_i\alpha(1-\alpha) \xrightarrow[\alpha\to+\infty]{} -\infty$

$$\frac{\partial L}{\partial x}(x,\nu,\lambda) = c + A^T\lambda - \nu + 2Diag(\nu)x = \mathbf{0}$$

Which gives:

$$g(\nu,\lambda) = \min_x L(x,\nu,\lambda) = \begin{cases} -\frac{1}{4}(c+A^T\lambda+\nu)^T(Diag(\nu))^{-1}(c+A^T\lambda+\nu) - \lambda^Tb \\ = -\frac{1}{4}\sum_{i=1}^{n}\frac{(c_i+a_i^T\lambda+\nu_i)^2}{\nu_i} - \lambda^Tb \quad (if\ \nu \succeq \mathbf{0}) \\ -\infty\ otherwise \end{cases}$$

Where a_i is the i^{th} column of A. If $\nu_i = 0$ then we set the term:

$$\frac{(c_i+a_i^T\lambda+\nu_i)^2}{\nu_i} = \begin{cases} +\infty, & if\ c_i+a_i^T\lambda+\nu_i \neq 0 \\ 0 & if\ c_i+a_i^T\lambda+\nu_i = 0 \end{cases}$$

The dual problem is:

$$\begin{aligned} \text{maximize} \quad & -\frac{1}{4}\sum_{i=1}^{n}\frac{(c_i + a_i^T\lambda + \nu_i)^2}{\nu_i} - \lambda^T b \\ \text{subject to} \quad & \nu \succeq \mathbf{0} \\ & \lambda \succeq \mathbf{0} \end{aligned}$$

We can then maximize over ν, knowing that:

$$\sup_{\nu_i \geq 0}\left(-\frac{1}{4}\frac{(c_i + a_i^T\lambda - \nu_i)^2}{v_i}\right) = \begin{cases} c_i + a_i^T\lambda \quad if\ c_i + a_i^T\lambda \leq 0 \\ 0 \quad if\ c_i + a_i^T\lambda > 0 \end{cases} = \min(0, c_i + a_i^T\lambda)$$

The equivalent problem is:

$$\begin{aligned} \text{maximize} \quad & -\lambda^T b - \sum_{i=1}^{n}\mu_i \\ \text{subject to} \quad & \mu_i = -\min(0, c_i + a_i^T\lambda) \\ & \lambda \succeq \mathbf{0} \end{aligned} \tag{13}$$

(b) We consider the LP relaxation problem:

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax \preceq b \\ & 0 \leq x_i \leq 1 :\ i = 1, .., n \end{aligned} \tag{14}$$

Which we can rewrite as:

$$\begin{aligned} \text{minimize} \quad & c^T x \\ \text{subject to} \quad & Ax \preceq b \\ & -x \leq 0 \\ & x \leq 1 \end{aligned}$$

The Lagrangian of the new problem is:

$$L(x, \lambda, \nu, \mu) = c^T x + \lambda^T(Ax - b) - \nu^T x + \mu^T(x - \mathbf{1}) = (c + A^T\lambda - \nu + \mu)^T x - \lambda^T b - \mu^T\mathbf{1}$$

Hence the dual function:

$$g(\lambda, \nu, \mu) = \begin{cases} -\lambda^T b - \mu^T\mathbf{1} \quad if\ c + A^T\lambda - \nu + \mu = \mathbf{0} \\ -\infty \quad otherwise \end{cases}$$

The explicit dual problem:

$$\begin{aligned} \text{maximize} \quad & -\lambda^T b - \mu^T \mathbf{1} \\ \text{subject to} \quad & c + A^T\lambda - \nu + \mu = \mathbf{0} \\ & \nu \succeq \mathbf{0} \\ & \mu \succeq \mathbf{0} \\ & \lambda \succeq \mathbf{0} \end{aligned} \tag{15}$$

Let $p^*_{relaxation}$ be the optimal value of the relaxation problem (14) and p^* that of the boolean problem (12). The feasible set of the relaxation includes the feasible set of the boolean, hence:

$$p^*_{relaxation} \leq p^*$$

On the other hand, the two Lagrangian problems (13) and (15) are equivalent when introducing $\nu = \mu + c + A^T\lambda$ in (13). Therefore, we end up with the same lower bound.