Convex Optimization Homework 2

Soël Megdoud

October 2024

1 Exercise 1

1)

To compute the dual, we introduce Lagrange multipliers $\mu \in R^n$ for the equality constraint $Ax = b$ and $\lambda \in R^d_+$ for the inequality constraint $x \geq 0$.

The Lagrangian is:

$$\begin{aligned} L(x, \mu, \lambda) &= c^\top x + \mu^\top (Ax - b) - \lambda^\top x \\ &= (c + A^\top \mu - \lambda)^\top x - \mu^\top b. \end{aligned}$$

To find the dual function $g(\mu, \lambda)$, we minimize $L(x, \mu, \lambda)$ over $x \geq 0$:

$$g(\mu, \lambda) = \inf_{x \geq 0} L(x, \mu, \lambda) = \begin{cases} -\mu^\top b & \text{if } c + A^\top \mu - \lambda = 0, \\ -\infty & \text{otherwise.} \end{cases}$$

Since $\lambda \geq 0$, the condition $c + A^\top \mu - \lambda = 0$ implies $c + A^\top \mu \geq 0$. Therefore, we can write the dual problem as:

$$\begin{aligned} \max_{\mu} \quad & -\mu^\top b \\ \text{s.t.} \quad & -A^\top \mu \leq c. \end{aligned}$$

Therefore, the dual of (P) is (D)

2)

The dual function is the following:

$$g(\mu, \lambda) = \inf_y (-b^\top y + \lambda^\top (A^\top y - c)) = \inf_y ((A\lambda - b)^\top y - c^\top \lambda) = \begin{cases} -c^\top \lambda & \text{if } A\lambda - b = 0, \\ -\infty & \text{otherwise.} \end{cases}$$

Therefore, the dual problem is:

$$\begin{aligned} \max_{\lambda} \quad & -\lambda^\top c \\ \text{s.t.} \quad & A\lambda = b, \lambda \geq 0 \end{aligned}$$

Which is the same as (P):

$$\begin{aligned} \min_{\lambda} \quad & \lambda^\top c \\ \text{s.t.} \quad & A\lambda = b, \lambda \geq 0 \end{aligned}$$

3)

The Lagrangian is given by:

$$\mathcal{L}(x, y, \mu, \lambda_1, \lambda_2) = c^T x - b^T y + \mu^T (b - Ax) + \lambda_1^T (A^T y - c) - \lambda_2^T x.$$

We can rewrite it:

$$\mathcal{L}(x, y, \mu, \lambda_1, \lambda_2) = (c - A^T \mu - \lambda_2)^T x + (A\lambda_1 - b)^\top y + b^\top \mu - c^\top \lambda_1$$

The dual function $g(\mu, \lambda_1, \lambda_2)$ is obtained by minimizing $\mathcal{L}$ over x and y. To be finite, we must have:

$$c - A^T\mu - \lambda_2 = 0 \implies c - A^T\mu = \lambda_2 \geq 0.$$

We then have the constraint:

$$c \geq A^T\mu$$

For the same reason, we also have this constraint:

$$A\lambda_1 - b = 0 \implies A\lambda_1 = b.$$

The dual problem is then :

$$\begin{aligned} \max_{\mu,\lambda_1} \quad & -c^\top\lambda_1 + b^\top\mu \\ \text{s.t.} \quad & A^T\mu \leq c, \\ & A\lambda_1 = b, \quad \lambda_1 \geq 0. \end{aligned}$$

This is problem is the same as the following, demonstrating that it is self-dual.
The dual problem is then :

$$\begin{aligned} \min_{\mu,\lambda_1} \quad & c^\top\lambda_1 - b^\top\mu \\ \text{s.t.} \quad & A^T\mu \leq c, \\ & A\lambda_1 = b, \quad \lambda_1 \geq 0. \end{aligned}$$

3)

a)

Solving the above problem boils down to solving both (P) and (D) and adding their results:

$$\begin{aligned} \min_{x,y} \quad & c^\top x - b^\top y \\ \text{s.t.} \quad & A^Ty \leq c, \\ & Ax = b, \quad \geq 0. \end{aligned}$$

This is equivalent to

$$\begin{aligned} & \min_x c^\top x + \max_y b^\top y \\ & \text{s.t.} \quad Ax = b, \quad x \geq 0 \quad \& \\ & \text{s.t.} \quad A^Ty \leq c. \end{aligned}$$

b)

Let x^* be an optimal solution to the primal problem (P):

$$\begin{aligned} \min_x \quad & c^\top x \\ \text{s.t.} \quad & Ax = b, x \geq 0. \end{aligned}$$

Let y^* be an optimal solution to the dual problem (D):

$$\begin{aligned} \max_y \quad & b^\top y \\ \text{s.t.} \quad & A^\top y \leq c. \end{aligned}$$

We showed in 4.a) that (x^*, y^*) is optimal for the Self-Dual problem.
Moreover, the strong duality theorem for linear programs states that optimal objective values of (P) and (D) are equal:

$$c^\top x^* = b^\top y^*.$$

We can now evaluate the objective function of the Self-Dual problem at (x^*, y^*):

$$\begin{aligned} c^\top x^* - b^\top y^* &= c^\top x^* - c^\top x^* \quad (\text{since } c^\top x^* = b^\top y^*) \\ &= 0. \end{aligned}$$

Therefore, the optimal value is zero.

2 Exercice 2

1)

The conjugate $f_1^*(y)$ is defined as:

$$f_1^*(y) = \sup_x \{y^\top x - \|x\|_1\}.$$

Expanding this for each component:

$$f_1^*(y) = \sup_x \left\{ \sum_{i=1}^{d} (y_i x_i - |x_i|) \right\}.$$

For each component:

- If $|y_i| \leq 1$, the supremum is finite and achievable, yielding a value of 0. We have:

  $$y_i x_i - |x_i| \leq |y_i x_i| - |x_i| \leq |x_i|(|y_i| - 1) \leq 0.$$

  and it is equal to zero for $x_i = 0$.

- If $y_i > 1$ or $y_i < -1$, the supremum becomes $+\infty$ as the function can be made unbounded.

Finaly we have :

$$f_1^*(y) = \begin{cases} 0, & \text{if } \|y\|_\infty \leq 1, \\ +\infty, & \text{otherwise.} \end{cases}$$

The conjugate of the ℓ_1 norm is:

$$f_1^*(y) = \delta_{\|y\|_\infty \leq 1}(y),$$

where $\delta_{\|y\|_\infty \leq 1}(y)$ is the indicator function that is 0 if $\|y\|_\infty \leq 1$ and $+\infty$ otherwise.

2)

Let's first compute the conjugate of the squared norm.
Let $f_2(x) = \|x\|_2^2$. The conjugate $f_2^*(y)$ is defined by:

$$f_2^*(y) = \sup_x \{y^\top x - \|x\|_2^2\}.$$

To find this, we take the gradient of $y^\top x - \|x\|_2^2$ with respect to x and set it to zero:

$$\nabla_x (y^\top x - \|x\|_2^2) = y - 2x = 0 \implies x = \frac{y}{2}.$$

Substitute $x = \frac{y}{2}$ back into the expression to find the supremum:

$$f_2^*(y) = y^\top \left(\frac{y}{2}\right) - \left\|\frac{y}{2}\right\|_2^2 = \frac{\|y\|_2^2}{2}.$$

Thus, the conjugate of the squared norm is:

$$f_2^*(y) = \frac{\|y\|_2^2}{4}.$$

Let's introduce $y = Ax - b$ to reformulate the problem to:

$$\min_{x,y} \|y\|_2^2 + \|x\|_1 \quad \text{s.t.} \quad y = Ax - b.$$

The Lagrangian $\mathcal{L}(x, y, \mu)$ is given by:

$$\mathcal{L}(x, y, \mu) = \|y\|_2^2 + \|x\|_1 + \mu^\top (Ax - b - y).$$

Separating the terms involving x and y:

$$\mathcal{L}(x, y, \mu) = (\mu^\top Ax + \|x\|_1) + (\|y\|_2^2 - \mu^\top y) - \mu^\top b.$$

To find the dual function $g(\mu)$, we take the infimum over x and y separately:

$$g(\mu) = \inf_x(\mu^\top Ax + \|x\|_1) + \inf_y(\|y\|_2^2 - \mu^\top y) - \mu^\top b.$$

We can rewrite

$$\inf_x(\mu^\top Ax + \|x\|_1) = -f_1^*(-A^\top \mu),$$

where $f_1^*(-A^\top \mu)$ is the conjugate of the ℓ_1 norm, as calculated previously:

$$f_1^*(-A^\top \mu) = \delta_{\|A^\top \mu\|_\infty \leq 1}(-A^\top \mu).$$

We do the same over y

$$\inf_y(\|y\|_2^2 - \mu^\top y) = -f_2^*(\mu) = -\frac{\|\mu\|_2^2}{4}.$$

Finally, we have:

$$g(\mu) = \delta_{\|A^\top \mu\|_\infty \leq 1}(-A^\top \mu) - \frac{\|\mu\|_2^2}{4} - \mu^\top b.$$

The dual problem is:

$$\max_\mu \quad -\mu^\top b - \frac{\|\mu\|_2^2}{4} \quad \text{s.t.} \quad \|A^\top \mu\|_\infty \leq 1.$$

3)

1.

We can write

$$\min_{\omega, z} \quad \frac{1}{n\tau} 1^\top z \quad \text{s.t.} \quad z_i \geq 1 - y_i(\omega^\top x_i), \quad z \geq 0.$$

As:

$$\min_{\omega, z} \quad \frac{1}{n\tau} \sum_{i=1}^n z_i \quad \text{s.t.} \quad z_i \geq 1 - y_i(\omega^\top x_i), \quad z \geq 0.$$

- When x_i is miss-classified $1 - y_i(\omega^\top x_i) > 0$ and $z_i \geq 0$ so $z_i = 1 - y_i(\omega^\top x_i)$.
- When x_i is well-classified $1 - y_i(\omega^\top x_i) \leq 0$ and $z_i \geq 0$ so $z_i = 0$

Therfore we can rewrite it

$$\min_\omega \quad \frac{1}{n\tau} \sum_{i=1}^n L(\omega, x_i, y_i)$$

where the loss function $L(\omega, x_i, y_i)$ is defined as:

$$L(\omega, x_i, y_i) = \max(0, 1 - y_i(\omega^\top x_i)).$$

Thus, we showed that solving (Sep. 2) solves (Sep. 1)

2.

To compute the dual of (Sep. 2), we start with the Lagrangian:

$$\mathcal{L}(\omega, z, \lambda, \pi) = \frac{1}{n\tau} 1^\top z + \frac{1}{2}\|\omega\|_2^2 + \sum_{i=1}^n \lambda_i(1 - y_i(\omega^\top x_i) - z_i) - \pi^\top z,$$

We can group terms involving ω and z

$$\mathcal{L}(\omega, z, \lambda, \pi) = \frac{1}{2}\|\omega\|_2^2 - \sum_{i=1}^n \lambda_i y_i(\omega^\top x_i) + \sum_{i=1}^n \left(\frac{1}{n\tau} - \lambda_i - \pi_i\right) z_i + \sum_{i=1}^n \lambda_i.$$

For the minimum over ω:

$$\inf_{\omega} \left(\frac{1}{2}\|\omega\|_2^2 - \sum_{i=1}^{n} \lambda_i y_i (\omega^\top x_i) \right).$$

The stationarity condition gives:

$$\omega = \sum_{i=1}^{n} \lambda_i y_i x_i.$$

Substitute ω back into the minimum over ω we get:

$$\inf_{\omega} \left(\frac{1}{2}\|\omega\|_2^2 - \sum_{i=1}^{n} \lambda_i y_i (\omega^\top x_i) \right) = \inf_{\omega} \left(\frac{1}{2}\|\omega\|_2^2 - |\omega\|_2^2 \right) = -\frac{1}{2} \left\| \sum_{i=1}^{n} \lambda_i y_i x_i \right\|_2^2.$$

For the infimum over z:

$$\inf_{z} \left(\sum_{i=1}^{n} \left(\frac{1}{n\tau} - \lambda_i - \pi_i \right) z_i \right).$$

For the dual to be finite, we require:

$$\frac{1}{n\tau} - \lambda_i - \pi_i = 0 \implies \pi_i = \frac{1}{n\tau} - \lambda_i, \quad \text{with } \pi_i \geq 0.$$

This implies:

$$0 \leq \lambda_i \leq \frac{1}{n\tau}.$$

The dual problem can be formulated as:

$$\max_{\lambda} \quad \sum_{i=1}^{n} \lambda_i - \frac{1}{2} \left\| \sum_{i=1}^{n} \lambda_i y_i x_i \right\|_2^2 \quad \text{s.t.} \quad 0 \leq \lambda_i \leq \frac{1}{n\tau}, \quad \forall i.$$