MATH60005/70005: Optimisation (Autumn 23-24)

Week 10: Problem Session

Dr Dante Kalise, Dr Estefanía Loayza-Romero and Sara Bicego
Department of Mathematics
Imperial College London, United Kingdom
{dkaliseb,kloayzar,s.bicego21}@imperial.ac.uk

1. Solve the primal and dual problem for

$$\begin{array}{ll} \min & x_1^2 + x_2^2 + 2x_1 \\ \text{s.t.} & x_1 + x_2 = 0\,. \end{array}$$

2. Study the duality gap (difference between f^* and q^*) for the problem

$$\min\left\{e^{-x_2} : \sqrt{x_1^2 + x_2^2} - x_1 \le 0\right\}\,.$$

3. Recompute the dual of the convex quadratic problem from the notes under that assumption that the matrix $\mathbf{Q} \succeq 0$ instead of $\mathbf{Q} \succ 0$.

4. Consider the Chebyshev center problem where we have a set of points $\mathbf{a}_1 \ldots, \mathbf{a}_m \in \mathbb{R}^n$ for which we seek a point $\mathbf{x} \in \mathbb{R}^n$ that is the center of a ball of minimum radius $r > 0$ containing the points

$$\begin{array}{ll} \min_{\mathbf{x},r} & r \\ \text{s.t.} & \|\mathbf{x} - \mathbf{a}_i\| \le r\,, \quad i = 1, 2, \ldots, m\,. \end{array}$$

Compute the dual of this problem. *(Hint: use an equivalent formulation over the squared radius)*

Solutions

1. The primal problem has the following quadratic objective

$$f(\mathbf{x}) = \mathbf{x}^\top \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \mathbf{x} + 2\begin{bmatrix} 1 & 0 \end{bmatrix} \mathbf{x}$$

which is convex (as the identity matrix is positive definite). Hence, KKT condition are in this case necessary and sufficient.
For the Lagrangian

$$\mathcal{L}(\mathbf{x}, \mu) = x_1^2 + x_2^2 - 2x_1 + \mu(x_1 + x_2)$$

we have the following KKT system:

$$\begin{cases} 2x_1 + 2 + \mu = 0 & \\ 2x_2 + \mu = 0 & (\mu = -2x_2) \\ x_1 + x_2 = 0 & (x_1 = -x_2) \end{cases} \implies (x_1^*, x_2^*) = \left(-\frac{1}{2}, \frac{1}{2}\right)$$

at which $f(\mathbf{x}^*) = -\frac{1}{2}$.

Since the objective function $f(\mathbf{x})$ is convex, and the linear equation is feasible, strong duality holds for this problem. For the associated Lagrangian $\mathcal{L}(\mathbf{x}, \mu)$ given above, we have

$$\mathbf{q}(\mu) = \min_{\mathbf{x} \in \mathbb{R}^2} \mathcal{L}(\mathbf{x}, \mu),$$

and

$$\nabla_{\mathbf{x}} \mathcal{L} = 0 \iff \begin{cases} 2x_1 + 2 + \mu = 0 \\ 2x_2 + \mu = 0 \end{cases} \implies \begin{cases} x_2 = -\frac{\mu}{2} \\ x_1 = -1 - \frac{\mu}{2} \end{cases} \qquad (*)$$

for which the dual objective becomes

$$\begin{aligned} \mathbf{q}(\mu) &= (-1 - \frac{\mu}{2})^2 + (-\frac{\mu}{2})^2 + 2(-1 - \frac{\mu}{2}) + \mu(-1 - \mu) \\ &= -\frac{\mu^2}{2} - \mu - 1 . \end{aligned}$$

Then, the dual problem reads

$$\max_{\mu} -\frac{\mu^2}{2} - \mu - 1$$

which is realized when $\mathbf{q}'(\mu) = 0$, hence at $\mu^* = -1$. Then, from $(*)$ the minimizer is

$$(x_1^*, x_2^*) = \left(-\frac{1}{2}, \frac{1}{2}\right)$$

as in the solution of the primal.

2. We consider

$$\min e^{-x_2} \quad \text{s.t.} \quad \sqrt{x_1^2 + x_2^2} - x_1 \leq 0$$

for which the feasible set is $\{\mathbf{x} : x_2 = 0,\ x_1 \geq 0\}$, hence Slater's condition cannot be satisfied, and so only week duality holds ($f^* > q^*$). Since optimility requires $x_2 = 0$, we have that for the primal, $f^* = 1$. Moving now to the dual problem, we have associated Lagrangian

$$\mathcal{L}(\mathbf{x}, \lambda) = e^{-x_2} + \lambda(\sqrt{x_1^2 + x_2^2} - x_1), \quad \lambda \geq 0,$$

for which the dual objective is $\mathbf{q}(\mu) = \min_{\mathbf{x}\in\mathbb{R}^2} \mathcal{L}(\mathbf{x}, \mu)$. Note that $\mathcal{L} \geq 0$, but for any $\varepsilon > 0$, we have

$$\mathbf{x}_\varepsilon = \left(\frac{x_{2,\varepsilon}^2 - \varepsilon^2}{2\varepsilon}, ln(\varepsilon)\right) \implies \sqrt{x_{1,\varepsilon}^2 + x_{2,\varepsilon}^2} - x_{1,\varepsilon} = \varepsilon \implies \mathcal{L}(\mathbf{x}_\varepsilon, \lambda) = (1+\lambda)\varepsilon$$

and so

$$\mathbf{q}(\mu) = \min_{\mathbf{x}\in\mathbb{R}^2} \mathcal{L}(\mathbf{x}, \mu) \iff \min_{\varepsilon>0} \mathcal{L}(\mathbf{x}_\varepsilon, \lambda) \iff \min_{\varepsilon>0}(1+\lambda)\varepsilon \longrightarrow 0 \quad \forall \lambda \geq 0\,.$$

Thus, $\mathbf{q}(\mu) = 0$ and the dual solution is given by

$$\max_{\lambda\geq 0} \left\{0\right\} = \mathbf{q}^* = 0\,.$$

3. The quadratic problem from the notes is

$$\min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{x}^\top Q\mathbf{x} + 2f^\top \mathbf{x}\,, \qquad \text{s.t.} \qquad A\mathbf{x} \leq b\,,$$

where we assume $Q \succeq 0$. This implies that Q is not necessarily invertible, but there exists a matrix $D \in \mathbb{R}^{n\times n}$ such that $Q = D^\top D$. Thus, the quadratic problem is equivalent to

$$\min_{\mathbf{x}\in\mathbb{R}^n} \mathbf{x}^\top D^\top D\mathbf{x} + 2f^\top \mathbf{x}\,, \qquad \text{s.t.} \qquad A\mathbf{x} \leq b\,,$$

which in turns is equivalent to

$$\min_{\mathbf{x},\mathbf{z}\in\mathbb{R}^n} ||\mathbf{z}||^2 + 2f^\top \mathbf{x}\,, \qquad \text{s.t.} \qquad \begin{cases} A\mathbf{x} \leq b \\ bz = D\mathbf{x} \end{cases}$$

The associated Lagrangian

$$\mathcal{L}(\mathbf{x}, \mathbf{z}, \lambda, \mu) = ||\mathbf{z}||^2 + 2f^\top \mathbf{x} + 2\lambda^\top(A\mathbf{x} - b) + 2\mu^\top(\mathbf{z} - D\mathbf{x})$$

it is separable with respect to $\mathbf{x}$ and $\mathbf{z}$, hence the objective for the dual problem can be computed as

$$\mathbf{q}(\lambda, \mu) = \underbrace{\min_{\mathbf{z}}\left(||\mathbf{z}||^2 + 2\mu^\top \mathbf{z}\right)}_{(a)} + 2\underbrace{\min_{\mathbf{x}}\left(f^\top \mathbf{x} + \lambda^\top A\mathbf{x} - \mu^\top D\mathbf{x}\right)}_{(b)} - 2\lambda^\top b$$

(a) by first order optimality conditions, we have

$$\nabla_{\mathbf{z}}\left(||\mathbf{z}||^2 + 2\mu^\top \mathbf{z}\right) = 0 \implies 2\mathbf{z} + 2\mu = 0 \implies \mathbf{z}^* = -\mu$$

(b) we can rearrange the objective and obtain

$$\min_{\mathbf{x}}\left((f^\top + \lambda^\top A - D^\top \mu)\mathbf{x}\right) = \begin{cases} 0 & f + A^\top\lambda - D^\top\mu = 0 \\ -\infty & \text{otherwise} \end{cases}$$

Thus, the dual objective reads

$$\mathbf{q}(\lambda, \mu) = \min_{\mathbf{x},\mathbf{q}} \mathcal{L}(\mathbf{x}, \mathbf{z}, \lambda, \mu) = \begin{cases} \|\mu\|^2 - 2\lambda^\top b & f + A^\top \lambda - D^\top \mu = 0 \\ -\infty & \text{otherwise} \end{cases}$$

and since $\{-\infty\} \not\subset Dom(\mathbf{q})$, we can write the dual problem as

$$\max_{\substack{\mu \in \mathbb{R}^n \\ \lambda \in \mathbb{R}^m_+}} \|\mu\|^2 - 2\lambda^\top b \quad \text{s.t.} \quad f + A^\top \lambda - D^\top \mu = 0\,.$$

4. We consider the set $\{\mathbf{a}_i\}_{i=1}^m \subset \mathbb{R}^n$ of data points encircled by a ball of radius r and center $\mathbf{x}$:

$$\min_{\mathbf{x},r} r \quad \text{s.t.} \quad \|\mathbf{x} - \mathbf{a}_i\| \leq r \qquad i = 1, \ldots, m\,,$$

which is equivalent to

$$\min_{\mathbf{x},\gamma} \gamma \quad \text{s.t.} \quad \|\mathbf{x} - \mathbf{a}_i\|^2 \leq \gamma \qquad i = 1, \ldots, m\,.$$

The associated Lagrangian is

$$\mathcal{L}(\mathbf{x}, \gamma, \lambda) = \gamma + \sum_{i=1}^m \lambda_i (\|\mathbf{x} - \mathbf{a}_i\|^2 - \gamma)\,, \qquad \lambda \geq 0\,,$$

for which the dual objective becomes

$$\mathbf{q}(\lambda) = \min_{(\mathbf{x},\gamma) \in X} \mathcal{L}(\mathbf{x}, \gamma, \lambda) = \underbrace{\gamma(1 - \sum_{i=1}^m \lambda_i)}_{0} + \sum_{i=1}^m \lambda_i (\|\mathbf{x} - \mathbf{a}_i\|^2)\,.$$

We take $X = \mathbb{R}^n+1$, without imposing any restriction on γ (which should be positive, as $\gamma = r^2$). Assuming $\lambda \in \Delta_m$, we have

$$\sum_{i=1}^m \lambda_i \|\mathbf{x} - \mathbf{a}_i\|^2 = \sum_{i=1}^m \lambda_i \big(\|\mathbf{x}\|^2 - 2\mathbf{a}_i^\top \mathbf{x} + \|\mathbf{a}_i\|^2\big) = \|\mathbf{x}\|^2 \underbrace{\big(\sum_{i=1}^m \lambda_i\big)}_{1} - 2\big(\sum_{i=1}^m \lambda_i \mathbf{a}_i^\top\big)\mathbf{x} + \big(\sum_{i=1}^m \lambda_i \|\mathbf{a}_i\|^2\big),$$

and so

$$\mathbf{q}(\lambda) = \min_{\mathbf{x}} \Big(\|\mathbf{x}\|^2 - 2\big(\sum_{i=1}^m \lambda_i \mathbf{a}_i^\top\big)\mathbf{x} + \big(\sum_{i=1}^m \lambda_i \|\mathbf{a}_i\|^2\big)\Big).$$

By first order optimality conditions, we have

$$\mathbf{x}^* = \sum_{i=1}^m \lambda_i \mathbf{a}_i^\top = A\lambda$$

for the matrix A having as $i-$th column the coordinate of the point $\mathbf{a}_i$, for $i = 1, \ldots, m$. Thus

$$\mathbf{q}(\lambda) = \|A\lambda\|^2 - 2\underbrace{(A\lambda)^\top (A\lambda)}_{\|A\lambda\|^2} + \sum_{i=1}^m \lambda_i \|\mathbf{a}_i\|^2 = \begin{cases} -\|A\lambda\|^2 + \sum_{i=1}^m \lambda_i \|\mathbf{a}_i\|^2 & \lambda \in \Delta_m \\ -\infty & \text{otherwise} \end{cases}$$

for which the dual problem reads

$$\max_{\lambda \geq 0} \left(-\|A\lambda\|^2 + \sum_{i=1}^{m} \lambda_i \|\mathbf{a}_i\|^2 \right) \qquad \text{s.t.} \qquad \lambda \in \Delta_m.$$