Optimization theory exercises

Pablo Miralles González

October 31, 2022

1. Consider the general nonlinear optimization problem

$$\begin{array}{lll} \min & f(x) & \\ \text{s.t.} & g_i(x) \leqslant 0 & i = 1, \ldots, m \\ & h_i(x) = 0 & i = 1, \ldots, p. \end{array}$$

   (a) Write down the Lagrangian.

   Solution.
   $$L(x, \lambda, \mu) = f(x) + \lambda^T h(x) + \mu^T g(x).$$

   (b) Write down the dual function.

   Solution.
   $$q(\lambda, \mu) = \inf_{x \in \mathbb{R}^n} L(x, \lambda, \mu).$$

   (c) Write down the dual problem.

   Solution.
   $$\begin{array}{ll} \max & q(\lambda, \mu) \\ \text{s.t.} & \mu \geqslant 0 \end{array}$$

   (d) Explain weak duality. Show that weak duality always holds.

   Solution. The term weak duality refers to the fact that the optimal cost of the dual problem is always a lower bound of the optimal cost of the primal problem. To show this, let x be any primal feasible point and (λ, μ) any dual feasible point. Then:
   $$q(\lambda, \mu) \leqslant L(x, \lambda, \mu) = f(x) + \lambda^T h(x) + \mu^T g(x) \underbrace{=}_{h(x)=0,\, \mu \geqslant 0,\, g(x) \leqslant 0} f(x).$$

   It suffices now to take infimums over dual and primal feasible points.

   (e) Explain what is strong duality. Suppose that all functions are convex. Give a sufficient condition under which strong duality holds.

   Solution. Strong duality refers to the condition $d^* = p^*$, where d^* is the optimal cost of the dual problem and p^* the optimal cost of the primal problem. This does not always hold, but when all functions are convex then Slater's condition is sufficient for strong duality to hold. In this case, each h_i should be affine, and that there should exist a feasible vector $\overline{x}$ such that $g_i(\overline{x}) < 0$ for each non-affine g_i.

2. Consider the following quadratic program

$$\begin{aligned} \min \quad & \frac{1}{2}x^T Bx + g^T x \\ \text{s.t.} \quad & Ax \leqslant b. \end{aligned}$$

(a) Under what condition is the cost a convex function? When is it strongly convex?

Solution. We know from class that $f(x) = \frac{1}{2}x^T Bx + g^T x$ is convex if H is positive semidefinite, whereas it is strongly convex if H is positive definite. To show this, we can use the fact that $\nabla^2 f(x) \equiv H$ and *Theorem 5.7*.

As a matter of fact, we can show that the reciprocal is also true. If H is not positive semidefinite, then there exists $\hat{x}$ such that $\hat{x}^T H\hat{x} < 0$. Then

$$f(\hat{x}) = \frac{1}{2}\hat{x}^T B\hat{x} + g^T\hat{x} < g^T\hat{x} = 0 + g^T(\hat{x} - 0) = f(0) + \nabla f(0)^T(\hat{x} - 0),$$

showing that the condition in *Theorem 5.6.(i)* does not hold. If suppose now that H is not positive definite, then there exists $\hat{x}$ such that $\hat{x}^T H\hat{x} \leqslant 0$, and by an analogous argument we arrive at

$$f(\hat{x}) \leqslant f(0) + \nabla f(0)^T(\hat{x} - 0),$$

which shows that the condition in *Theorem 5.6.(iii)* cannot hold.

(b) Under what conditions is the problem convex?

Solution. Since $\{x \in \mathbb{R}^n : Ax \leqslant b\}$ is always a convex set, the problem is convex if and only if f is convex, that is, if H is positive semidefinite.

(c) Suppose that the cost is strongly convex. Derive the dual problem.

Solution. The Lagrangian and its derivatives with respect to x are:

$$\begin{aligned} L(x, \mu) &= \frac{1}{2}x^T Bx + g^T x + \mu^T(Ax - b) \\ \nabla_x L(x, \mu) &= Bx + g + A\mu \\ \nabla_x^2 L(x, \mu) &= B. \end{aligned}$$

Since f is strongly convex, B is positive definite and invertible. We deduce then that the only stationary point

$$x^* = -B^{-1}(g + A\mu)$$

is the only minimum. The dual function is therefore

$$q(\mu) = L(x^*, \mu) = -\frac{1}{2}\mu^T AB^{-1}A^T\mu - \mu^T(AB^{-1}g + b) - \frac{1}{2}g^T B^{-1}g.$$

The dual problem is thus

$$\begin{aligned} \max \quad & -\frac{1}{2}\mu^T AB^{-1}A^T\mu - \mu^T(AB^{-1}g + b) - \frac{1}{2}g^T B^{-1}g \\ \text{s.t.} \quad & \mu \geqslant 0. \end{aligned}$$

(d) When does strong duality hold?

Solution. We know that if Slater's conditions hold, then strong duality holds as well. In this case, since all constraints are affine, it suffices to show that f is convex, which happens if and only if H is positive semidefinite.

(e) Suppose that the cost is strongly convex, and that you apply an algorithm that solves the dual problem. Give a formula that recovers teh primal solution from a dual solution.

Solution. If μ^* is the solution of the dual problem, then by *Theorem 6.5*

$$x^* = arg \min_{x \in \mathbb{R}^n} L(x, \mu^*) = -B^{-1}(g + A\mu).$$

3. The (Euclidean) projection of a vector $z \in \mathbb{R}^n$ onto a subset C of $\mathbb{R}^n$ is defined as the solution of the following optimization problem:

$$\begin{aligned} \min \quad & \frac{1}{2}\|x - z\|^2 \\ s.t. \quad & x \in C \end{aligned} \tag{1}$$

and is denoted by $P_C(z)$.

(a) Assume that the set C is a convex set. Is the projection $P_C(z)$ unique? Why?

Solution. Yes, the cost function is a quadratic function with the identity matrix, which is positive definite. This means that the cost is strongly convex, and therefore there exists at most one global minimum (see *Theorem 4.2.*).

(b) Suppose that we want to compute $P_C(z)$ for $C = \{x \in \mathbb{R}^n | Ax = 0\}$, where A is an $m \times n$ matrix (set C is called the nullspace of A). In other words, $P_C(z)$ is the solution of

$$\begin{aligned} \min \quad & \frac{1}{2}\|x - z\|^2 \\ s.t. \quad & Ax = 0. \end{aligned} \tag{2}$$

Show that the dual function related to problem Equation 2 can be expressed as

$$q(\lambda) = -\frac{1}{2}\|A^T\lambda - z\|^2 + \frac{1}{2}\|z\|^2,$$

where $\lambda \in \mathbb{R}^m$ is the dual vector associated with the equality constraints of Equation 2. Show that the (negative) dual problem can be expressed as follows:

$$\begin{aligned} \min \quad & \frac{1}{2}\|y - z\|^2 \\ s.t. \quad & y \in C^\perp, \end{aligned} \tag{3}$$

where $C^\perp = \{y \in \mathbb{R}^m | y = A^T\lambda, \lambda \in \mathbb{R}^m\}$.

Solution. We have that

$$\begin{aligned} L(x, \lambda) &= \frac{1}{2}\|x - z\|^2 + \lambda^T Ax \\ \nabla_x L(x, \lambda) &= x - z + A\lambda \\ \nabla_x^2 L(x, \lambda) &= \mathrm{Id}_n . \end{aligned}$$

The only minimum is $x = z - A\lambda$, and therefore

$$q(\lambda) = \frac{1}{2}\|A\lambda\|^2 + \lambda^T A(z - A\lambda) = -\frac{1}{2}\|A\lambda\|^2 + \lambda^T Az \left(-\frac{1}{2}\|z\|^2 + \frac{1}{2}\|z\|^2\right) = -\frac{1}{2}\|A^T\lambda - z\|^2 + \frac{1}{2}\|z\|^2.$$

The last part follows from the fact that $\|z\|^2$ does not depend on λ, so we can remove the term from the dual problem.

(c) Explain why strong duality always hold for Equation 2. Use strong duality to show

$$\|z\|^2 = \|z - P_C(z)\|^2 + \|z - P_{C^\perp}(z)\|^2. \tag{4}$$

Solution. Slater's conditions hold, show strong duality holds as well. Let x^* be the solution of Equation 2 and λ^* the solution of the dual. Then

$$\frac{1}{2}\|x^* - z\|^2 = p^* = d^* = q(\lambda^*) = -\frac{1}{2}\|A^T\lambda^* - z\|^2 + \frac{1}{2}\|z\|^2.$$

By putting $y^* = A^T\lambda^*$, we have that

$$\|z\|^2 = \|x^* - z\|^2 + \|y^* - z\|^2,$$

but $x^* = P_C(z)$ and by the equivalence between the dual problem and Equation 3 we have that $y^* = P_{C^\perp}(z)$.

(d) Use Lagrangian stationarity to show that

$$z = P_C(z) + P_{C^\perp}(z).$$

Solution.

$$0 = \nabla_x L(x^*, \lambda^*) = x^* - z + A\lambda^* = x^* - z + y^*,$$

and therefore $z = P_C(z) + P_{C^\perp}(z)$.

(e) Suppose that $A = \begin{bmatrix} 1 & 0 \end{bmatrix}$ and $z = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Find $P_C(z), P_{C^\perp}(z)$. Sketch graphically sets $C, C^\perp$, points $z, P_C(z), P_{C^\perp}(z)$. Based on this figure give a geometric interpretation of condition 4.

Solution. $C = \text{span}(0,1)$ (y-axis) and $C^\perp = \text{span}(1,0)$ (x- axis), and so $P_C(z) = (0,1)$ and $P_{C^\perp}(z) = (1,0)$. What we are doing here is decomposing the vector along the x- and y- axis. The equation is the Pythagoras' theorem.