Optimization Methods - Duality

Bolei Zhang

December 1, 2024

1 Solutions

Problem 1: Get the Lagrange function and dual function of the following problem:

$$\begin{aligned} &\min\ x^T x \\ &\text{s.t. } Ax = b \end{aligned} \tag{1}$$

Solution: This problem has no inequality constraints and p (linear) equality constraints. The Lagrangian is $L(x, \nu) = x^T x + \nu^T (Ax - b)$, with domain $\mathbb{R}^n \times \mathbb{R}^p$. The dual function is given by $g(\nu) = \inf_x L(x, \nu)$. Since $L(x, \nu)$ is a convex quadratic function of x, we can find the minimizing x from the optimality condition

$$\nabla_x L(x, \nu) = 2x + A^T \nu = 0,$$

which yields $x = -(1/2)A^T \nu$. Therefore, the dual function is

$$g(\nu) = L(-(1/2)A^T \nu, \nu) = -(1/4)\nu^T A A^T \nu - b^T \nu,$$

which is a concave quadratic function, with domain $\mathbb{R}^p$. The lower bound property states that for any $\nu \in \mathbb{R}^p$, we have

$$-(1/4)\nu^T A A^T \nu - b^T \nu \leq \inf\{x^T x | Ax = b\}$$

.

Problem 2: Get the Lagrange function and dual function of the following problem:

$$\begin{aligned} &\min\ c^T x \\ &\text{s.t. } Ax = b, x \geq 0. \end{aligned} \tag{2}$$

Solution: To form the Lagrangian we introduce multipliers λ_i for the n inequality constraints and multiplier ν_i for the equality constraints, and obtain

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

The dual function is

$$g(\lambda, \nu) = \inf L(x, \lambda, \nu) = -b^T \nu + \inf (c + A^T \nu - \lambda)^T$$

which is easily determined analytically, since a linear function is bounded below only when it is identically zero. Thus, $g(\lambda, \nu) = -\infty$ except when $c + A^T \nu - \lambda = 0$, in which case it is $-b^T \nu$:

$$g(\lambda, \nu) = \begin{cases} -b^T \nu & A^T \nu - \lambda + c = 0, \\ -\infty & otherwise. \end{cases} \tag{3}$$

Note that the dual function g is finite only on a proper affine subset of $\mathbb{R}^m \times \mathbb{R}^p$. We will see that this is a common occurrence. The lower bound property (5.2) is nontrivial only when λ and ν satisfy $\lambda \geq 0$ and $A^T \nu - \lambda + c = 0$. When this occurs, $-b^T \nu$ is a lower bound on the optimal value of the LP.

Problem 3: Get the optimal value of the following problem via KKT conditions:

$$\begin{aligned} &\min\ (1/2)x^T Px + q^T x + r \\ &\text{s.t. } Ax = b, P \in S_+^n \end{aligned} \tag{4}$$

Solution: The KKT conditions for this problem are

$$Ax^* = b, Px^* + q + A^T \nu^* = 0,$$

which we can write as

$$\begin{bmatrix} P & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} x^* \\ \nu^* \end{bmatrix} = \begin{bmatrix} -q \\ b \end{bmatrix}$$

Solving this set of $m + n$ equations in the $m + n$ variables x^*, ν^* gives the optimal primal and dual variables.

Problem 4: Get the optimal value of the max entropy problem via KKT conditions:

$$\begin{aligned} &\min \sum_{i=1}^n x_i \log x_i \\ &\text{s.t. } Ax \leq b, \mathbf{1}^T x = 1 \end{aligned} \tag{5}$$

Solution: The Lagrange function is

$$L(x, \lambda, \nu) = \sum_{i=1}^n x_i \log x_i + \lambda^T (Ax - b) + \nu(\mathbf{1}^T x - 1)$$

By taking the derivative and setting it to zero:

$$\frac{\partial L}{\partial x_i} = \log x_i + 1 + a_i^T \lambda + \nu = 0$$

Therefore, $x_i = e^{-1-a_i^T \lambda - \nu}$. Substitute x_i in $g(\lambda, \nu)$

$$\begin{aligned} g(\lambda, \nu) &= \sum_{i=1}^n (x_i(-\lambda^T a_i - \nu - 1) + \lambda^T (a_i^T x_i - b_i) + \nu x_i) - \nu \\ &= -b^T \lambda - \nu - e^{-\nu - 1} \sum_{i=1}^n e^{-a_i^T \lambda} \end{aligned} \tag{6}$$

The dual problem is

$$\begin{aligned} &\max\ -b^T \lambda - \nu - e^{-\nu-1} \sum_{i=1}^n e^{-a_i^T \lambda} \\ &\text{s.t. } \lambda \geq 0 \end{aligned} \tag{7}$$

where a_i are the columns of A. We assume that the weak form of Slater's condition holds, i.e., there exists an $x > 0$ with $Ax \leq b$ and $\mathbf{1}^T x = 1$, so strong duality holds and an optimal solution (λ^*, ν^*) exists. Suppose we have solved the dual problem. The Lagrangian at (λ^*, ν^*) is

$$L(x, \lambda^*, \nu^*) = \sum_{i=1}^n x_i \log x_i + \lambda^{*T} (Ax - b) + \nu^* (\mathbf{1}^T x - 1)$$

which is strictly convex on D and bounded below, so it has a unique solution x^*, given by

$$x_i^* = 1/\exp(a_i^T \lambda^* + \nu^* + 1),\ i = 1, ..., n.$$

If x^* is primal feasible, it must be the optimal solution of the primal problem . If x^* is not primal feasible, then we can conclude that the primal optimum is not attained.

Problem 5: Get the optimal value of the water filling problem via KKT conditions:

$$\begin{aligned} &\min \ -\sum_{i=1}^{n} \log(\alpha_i + x_i) \\ &\text{s.t. } x \geq 0, \mathbf{1}^T x = 1 \end{aligned} \tag{8}$$

Solution: This problem arises in information theory, in allocating power to a set of n communication channels. The variable x_i represents the transmitter power allocated to the ith channel, and $\log(\alpha_i+x_i)$ gives the capacity or communication rate of the channel, so the problem is to allocate a total power of one to the channels, in order to maximize the total communication rate.

Introducing Lagrange multipliers $\lambda^* \in \mathbb{R}^n$ for the inequality constraints $x^* \geq 0$, and a multiplier $\nu^* \in \mathbb{R}$ for the equality constraint $\mathbf{1}^T x = 1$, we obtain the KKT conditions

$$x^* \geq 0, \mathbf{1}^T x^* = 1, \lambda^* \geq 0, \lambda_i^* x_i^* = 0, i = 1, ..., n, \ -1/(\alpha_i + x_i^*) - \lambda_i^* + \nu^* = 0, i = 1, ..., n.$$

We can directly solve these equations to find x^*, λ^*, and ν^*. We start by noting that λ^* acts as a slack variable in the last equation, so it can be eliminated, leaving

$$x^* \geq 0, \mathbf{1}^T x^* = 1, x_i^*(\nu^* - 1/(\alpha_i + x_i^*)) = 0, i = 1, ..., n, \nu^* \geq 1/(\alpha_i + x_i^*), i = 1, ..., n.$$

If $\nu^* < 1/\alpha_i$, this last condition can only hold if $x_i^* > 0$, which by the third condition implies that $\nu^* = 1/(\alpha_i + x_i^*)$. Solving for x_i^* , we conclude that $x_i^* = 1/\nu^* - \alpha_i$ if $\nu^* < 1/\alpha i$. If $\nu^* \geq 1/\alpha_i$, then $x_i^* > 0$ is impossible, because it would imply $\nu^* \geq 1/\alpha_i > 1/(\alpha_i + x_i^*)$, which violates the complementary slackness condition. Therefore, $x_i^* = 0$ if $\nu^* \geq 1/\alpha_i$. Thus we have

$$x_i^* = \begin{cases} 1/\nu^* - \alpha_i & \nu^* < 1/\alpha_i, \\ 0 & \textit{otherwise}. \end{cases} \tag{9}$$

or, put more simply, $x_i^* = \max\{0, 1/\nu^* - \alpha_i\}$. Substituting this expression for x_i^* into the condition $\mathbf{1}^T x^* = 1$ we obtain

$$\sum_{i=1}^{n} \max\{0, 1/\nu^* - \alpha_i\} = 1.$$

The lefthand side is a piecewise-linear increasing function of $1/\nu^*$, with breakpoints at α_i, so the equation has a unique solution which is readily determined.

This solution method is called water-filling for the following reason. We think of α_i as the ground level above patch i, and then flood the region with water to a depth $1/\nu$. The total amount of water used is then $\sum_{i=1}^{n} \max\{0, 1/\nu^* - \alpha - i\}$. We then increase the flood level until we have used a total amount of water equal to one. The depth of water above patch i is then the optimal value x_i^* .

Problem 6: Get the optimal value of the following problem via KKT conditions:

$$\begin{aligned} &\min \ x_1 \\ &\text{s.t. } 16 - (x_1 - 4)^2 - x_2^2 \geq 0 \\ &\quad\ \ x_1^2 + (x_2 - 2)^2 - 4 = 0 \end{aligned} \tag{10}$$

Solution: The Lagrange function is

$$L(x, \lambda, \nu) = x_1 + \nu(x_1^2 + (x_2 - 2)^2 - 4) - \lambda(16 - (x_1 - 4)^2 - x_2^2),$$

Get the stationary condition:

$$1 + 2\nu x_1 + 2\lambda(x_1 - 4) = 0, 2\nu(x_2 - 2) + 2\lambda x_2 = 0.$$

Consider the complementary slackness $\lambda = 0$ or $16 - (x_1 - 4)^2 - x_2^2 = 0$. If $\lambda = 0$, the stationary becomes

$$1 + 2\nu x_1 = 0, 2\nu(x_2 - 2) = 0. \tag{11}$$

From the first line, we can get $\nu \neq 0$, and therefore $x_2 = 2$. Taking the equality constraint that:

$$x_1^2 + (x_2 - 2)^2 - 4 = 0,$$

We can get $x1 = 2$. The KKT point is $(x_1, x_2, \lambda, \nu) = (2, 2, 0, -1/4)$.

If $16 - (x_1 - 4)^2 - x_2^2 = 0$ holds for the complementary slackness, combining $x_1^2 + (x_2 - 2)^2 - 4 = 0$, we can get two solutions $(0, 0)$ and $(8, 16)$ for (x_1, x_2). Therefore, the KKT points can be $(0, 0, 1/8, 0)$ and $(8/5, 16/5, 3/40, -1/5)$.

Considering the above KKT points, the optimal solution is $(0, 0, 1/8, 0)$.

Problem 7: Get the optimal value of the following problem via KKT conditions:

$$\begin{aligned} \min_{x \in \mathbb{R}, y > 0} \ & e^{-x} \\ \text{s.t.} \ & \frac{x^2}{y} \leq 0 \end{aligned} \tag{12}$$

Solution: According to the constraints, it is easy to convert the origin problem as

$$\begin{aligned} \min_{x \in \mathbb{R}} \ & e^{-x} \\ \text{s.t.} \ & x = 0 \end{aligned} \tag{13}$$

The Lagrange function is then

$$L(x, \nu) = e^{-x} + \nu x$$

And the dual function is

$$g(\nu) = \inf_x (e^{-x} + \nu x) = \begin{cases} \nu - \nu \ln \nu & \nu > 0, \\ 0, & \nu = 0, \\ -\infty & otherwise. \end{cases}$$

Therefore, the dual problem is

$$\max_\nu \begin{cases} \nu - \nu \ln \nu & \nu > 0, \\ 0, & \nu = 0, \end{cases}$$

The optimal solution is achieved when $\nu = 1$. The duality gap is 0.

Problem 8: Get the dual problem of

$$\begin{aligned} \min \ & x^T W x \\ \text{s.t.} \ & x_i^2 = 1, i = 1, ..., n \end{aligned} \tag{14}$$

Solution: The Lagrange function is

$$L(x, \nu) = x^T W x + \sum_{i=1}^n \nu_i (x_i^2 - 1) = x^T (W + \text{diag}(\nu)) x - 1^T \nu$$

The dual function is

$$\begin{aligned} g(\nu) &= \inf_x x^T (W + \text{diag}(\nu)) x - 1^T \nu \\ &= \begin{cases} -1^T \nu & W + \text{diag}(\nu) \succeq 0, \\ -\infty & otherwise. \end{cases} \end{aligned} \tag{15}$$

where we use the fact that the infimum of a quadratic form is either zero (if the form is positive semidefinite) or $-\inf$ (if the form is not positive semidefinite).

This dual function provides lower bounds on the optimal value. For example, we can take the specific value of the dual variable $\nu = -\lambda_{\min}(W) 1$, which is dual feasible, since $W + \text{diag}(\nu) = W - \lambda_{\min}(W) I \succeq 0$. This yields the bound on the optimal value p^*: $p^* \geq -1^T \nu = n \lambda_{\min}(W)$.

Problem 9: Get the conjugate function of

$$\begin{aligned} \min \ & f(x) \\ \text{s.t.} \ & x = 0 \end{aligned} \tag{16}$$

Solution: This problem has Lagrangian $L(x, \nu) = f(x) + \nu^T x$, and dual function

$$g(\nu) = \inf_x (f(x) + \nu^T x) = -\sup_x ((-\nu)^T x - f(x)) = -f^*(-\nu).$$

.

Problem 10: Get the dual problem of

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & Ax \leq b, Cx = d \end{aligned} \tag{17}$$

Solution: Using the conjugate of f_0 we can write the dual function for the problem as

$$\begin{aligned} g(\lambda, \nu) &= \inf_x (f_0(x) + \lambda^T (Ax - b) + \nu^T (Cx - d)) \\ &= -b^T \lambda - d^T \nu + \inf_x (f_0(x) + (A^T \lambda + C^T \nu)^T x) \\ &= -b^T \lambda - d^T \nu - f_0^*(-A^T \lambda - C^T \nu). \end{aligned} \tag{18}$$

The domain of g followsin from the domain of f_0^*:

$$\text{dom} g = \{(\lambda, \nu) | -A^T \lambda - C^T \nu \in \text{dom} f_0^*\}.$$