Introduction to Machine Learning - HW2 - Q6

Professors: Abolghasemi & Arabi
Student: Mohamad Mahdi Samadi
Student ID: 810101465

Consider this optimization problem:

$$\min_{x \in R^2} \; f(x) = \frac{1}{2}x^T Q x \; + c^T x \; subject \; to \; Ax \leq \; b$$

Where:

- $x \in R^2$: Decision variable vector
- $Q \in R^{2\times 2}$: 2X2 matrix that ensures a unique minimum.
- $c \in R^2$: 2-d vector.
- $A \in R^{3\times 2}$: 2X3 matrix representing the coefficients of the constraint.
- $b \in R^3$: 3-d vector representing the RHS of the constraint.

Find the dual problem in terms of the Lagrangian multipliers.

We have an optimization problem with a quadratic objective function and linear inequality constraints. The goal is to minimize the objective function subject to these constraints.

The Lagrangian function combines the objective function and the constraints using multipliers (Lagrange multipliers).

$$L(x, \lambda) = \frac{1}{2}x^T Q x + c^T x + \lambda^T (Ax - b)$$

Where λ is a 3-d vector of Lagrange multipliers, one for each constraint.

The KKT conditions are necessary for optimality in constrained optimization problems. For our problem, the conditions include:

- **Stationarity**: The gradient of the Lagrangian with respect to x should be equal to zero.

$$\nabla L(x,\lambda) = \frac{\partial(L(x,\lambda)}{\partial(x)} = \frac{1}{2}\left[\frac{\partial(x^TQx)}{\partial(x)}\right] + \frac{\partial(c^Tx)}{\partial(x)} + \frac{\partial(\lambda^T(Ax-b))}{\partial(x)}$$

$$\nabla L(x,\lambda) = \frac{1}{2}(Q + Q^T)x + c + A^T\lambda$$

$$\nabla L(x,\lambda) = Qx + c + A^T\lambda = 0 \rightarrow x = -Q^{-1}(c + A^T\lambda)$$

- **Dual Feasibility**: The Lagrange multipliers must be non-negative.

  $$\lambda \geq 0$$

- **Complementary Slackness**: For each constraint, the product of the Lagrange multiplier and the constraint must be zero.

  $$\forall i \rightarrow \lambda_i(A_i x - b_i) = 0$$

Now that we've checked the KKT conditions, let's continue to solve the equation.
We just need to replace the x back into the Lagrangian equation:

$$L(-Q^{-1}(c + A^T\lambda),\lambda) = \frac{1}{2}(-Q^{-1}(c + A^T\lambda))^T Q(-Q^{-1}(c + A^T\lambda)) +$$

$$+ c^T(-Q^{-1}(c + A^T\lambda)) + \lambda^T(A(-Q^{-1}(c + A^T\lambda)) - b)$$

$$L(x,\lambda) = \frac{1}{2}c^TQ^{-1}c + \frac{1}{2}c^TQ^{-1}A^T\lambda + \frac{1}{2}\lambda^TAQ^{-1}c + \frac{1}{2}\lambda^TAQ^{-1}A^T\lambda$$

$$- c^TQ^{-1}c - c^TQ^{-1}A^T\lambda - \lambda^TAQ^{-1}c - \lambda^TAQ^{-1}A^T\lambda - \lambda^TAb$$

Simplifying this expression results the dual function (let's call it $g(\lambda)$)

$$g(\lambda) = \frac{-1}{2}(c + A^T\lambda)^TQ^{-1}(c + A^T\lambda) - \lambda^Tb$$

The dual problem is to find maximum of the $g(\lambda)$ for all λ in the KKT conditions.

Dual Feasibility indicates that only non-negative λ values are acceptable.

Dual Problem: $\max_{\forall\lambda\geq 0} g(\lambda)$