Duality, Sensitivity, KKT

Computational Intelligence, Lecture 13

by Sergei Savin

Spring 2025

Content

- Lagrange dual function
- Duality gap, strong and weak duality
- KKT conditions
- Sensitivity

Consider an optimization problem:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f_0(\mathbf{x}), \\ \text{subject to} \quad & \begin{cases} f_i(\mathbf{x}) \le 0, \\ h_j(\mathbf{x}) = 0. \end{cases} \end{aligned} \tag{1}$$

It's *Lagrangian* is given as:

$$L(\mathbf{x}, \lambda_i, \nu_j) = f_0(\mathbf{x}) + \sum_i \lambda_i f_i(\mathbf{x}) + \sum_j \nu_j h_j(\mathbf{x}) \tag{2}$$

where λ_i and ν_j are Lagrange multipliers; they are sometimes called *dual variables*.

Given *Lagrangian* $L(\mathbf{x}, \lambda_i, \nu_j) = f_0(\mathbf{x}) + \sum_i \lambda_i f_i(\mathbf{x}) + \sum_j \nu_j h_j(\mathbf{x})$, the associated *Lagrange dual function* is given as:

$$g(\lambda_i, \nu_j) = \inf_{\mathbf{x}} L(\mathbf{x}, \lambda_i, \nu_j). \tag{3}$$

Lagrange dual function is **always concave**. If p^* is the optimal value of the cost function of the original problem, then $g(\lambda_i, \nu_j)$ gives as a *lower bound* on its possible values.

In fact, substituting any ν_j and $\lambda_i > 0$ gives us a valid lower bound on the cost. Maximum of $g(\lambda_i, \nu_j)$ over the domain given by $\lambda_i > 0$ provides us optimal (largest) lower bound of the problem, denoted as g^*.

Duality gap, strong and weak duality

If p^* is the optimal value of the cost function of the original problem and g^* is the optimal lower bound of the problem, then $p^* - g^*$ is called optimal *duality gap*.

If optimal duality gap is zero, the problem is said to have *strong duality*. If optimal duality gap greater than zero, the problem is said to have *weak duality*.

Lagrange dual function for a QP, 1

Consider the following QP:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & \mathbf{x}^\top \mathbf{H} \mathbf{x}, \\ \text{subject to} \quad & \mathbf{A}\mathbf{x} \leq \mathbf{b}. \end{aligned} \tag{4}$$

Its Lagrangian is:

$$L(\mathbf{x}, \lambda) = \mathbf{x}^\top \mathbf{H} \mathbf{x} + \lambda^\top (\mathbf{A}\mathbf{x} - \mathbf{b}) \tag{5}$$

In order to minimize the Lagrangian with respect to $\mathbf{x}$ we find the gradient and set it to zero:

$$\frac{\partial L(\mathbf{x}, \lambda)}{\partial \mathbf{x}} = 2\mathbf{x}^\top \mathbf{H} + \lambda^\top \mathbf{A} = 0 \tag{6}$$

With that we can compute $\mathbf{x}$ as a function of λ:

$$\mathbf{x} = -0.5\mathbf{H}^{-1}\mathbf{A}^\top \lambda \tag{7}$$

Knowing that $\mathbf{x} = -0.5\mathbf{H}^{-1}\mathbf{A}^\top \lambda$ we can compute $g(\lambda)$ by substituting the $\mathbf{x}$ we found into the Lagrangian:

$$g(\lambda) = \frac{1}{4}\lambda^\top \mathbf{A}\mathbf{H}^{-1}\mathbf{H}\mathbf{H}^{-1}\mathbf{A}^\top \lambda - \frac{1}{2}\lambda^\top \mathbf{A}\mathbf{H}^{-1}\mathbf{A}^\top \lambda - \lambda^\top \mathbf{b} \tag{8}$$

$$g(\lambda) = -\frac{1}{4}\lambda^\top \mathbf{A}\mathbf{H}^{-1}\mathbf{A}^\top \lambda - \lambda^\top \mathbf{b} \tag{9}$$

In order to find the optimal lower bound we solve the following problem:

$$\begin{aligned} \underset{\lambda}{\text{maximize}} \quad & -\frac{1}{4}\lambda^\top \mathbf{A}\mathbf{H}^{-1}\mathbf{A}^\top \lambda - \lambda^\top \mathbf{b}, \\ \text{subject to} \quad & \lambda \geq 0. \end{aligned} \tag{10}$$

Solution of this problem is the solution of the original problem.

Karush-Kuhn-Tucker (KKT) conditions certify optimality of an optimization problem:

$$\begin{aligned} &\underset{\mathbf{x}}{\text{minimize}} && f_0(\mathbf{x}), \\ &\text{subject to} && \begin{cases} f_i(\mathbf{x}) \leq 0, \\ h_j(\mathbf{x}) = 0. \end{cases} \end{aligned} \tag{11}$$

1. Primal feasibility: $f_i(\mathbf{x}) \leq 0$ and $h_j(\mathbf{x}) = 0$.
2. Dual feasibility: $\lambda_i \geq 0$.
3. Complementarity slackness: $\lambda_i f_i(\mathbf{x}) = 0$.
4. Lagrangian stationarity:
$\frac{\partial}{\partial \mathbf{x}} \left(f_0(\mathbf{x}) + \sum_i \lambda_i f_i(\mathbf{x}) + \sum_j \nu_j h_j(\mathbf{x}) \right) = 0.$

Optimal values of Lagrange variables λ determine local sensitivity of the system with respect to small perturbations of constraints.

Consider perturbed optimization problem:

$$\begin{aligned} \underset{\mathbf{x}}{\text{minimize}} \quad & f_0(\mathbf{x}), \\ \text{subject to} \quad & \begin{cases} f_i(\mathbf{x}) \le u_i, \\ h_j(\mathbf{x}) = v_j. \end{cases} \end{aligned} \tag{12}$$

where u_i and v_j - perturbations of the constraints. Let $p(\mathbf{u}, \mathbf{v})$ be optimal value of the cost function for given values of u_i and v_j. Then $p(0, 0)$ is the optimal value of the unperturbed problem.

Sensitivity, 2

Sensitivity of the optimal value of the cost function to the constraint perturbation is given as:

$$\left.\frac{\partial p(\mathbf{u}, \mathbf{v})}{\partial \mathbf{u}}\right|_{\mathbf{u}=0, \mathbf{v}=0} = -\lambda^*; \tag{13}$$

$$\left.\frac{\partial p(\mathbf{u}, \mathbf{v})}{\partial \mathbf{v}}\right|_{\mathbf{u}=0, \mathbf{v}=0} = -\nu^*. \tag{14}$$

Thus, analysing values of lagrange variables allows us to assess local sensitivity to constraint perturbation.

- Convex Optimization, Lecture 12: KKT conditions, Ryan Tibshirani.
- EE 227A: Convex Optimization and Applications, Lecture 13: Optimality Conditions for Convex Problems, Laurent El Ghaoui.
- The Karush–Kuhn–Tucker (KKT) Conditions (video). Visually Explained.

Lecture slides are available via Github, links are on Moodle:

github.com/SergeiSa/Computational-Intelligence-2025