Duality

Mathématiques Informatique et Statistique Appliquées (MISA)
Université d'Antananarivo

Lagrangian

Problem in standard form (not necessarily convex)

$$\begin{aligned} \text{minimize } & f_0(x) \\ \text{s.t } & f_i(x) \leq 0, \; i = 1, \ldots, m \\ & h_i(x) = 0, \; i = 1, \ldots, p \end{aligned}$$

where variable is $x \in \mathbb{R}^n$, domain $\mathcal{D}$, optimal value $p^\star$
The **Lagrangian** $L : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ associated with the problem is:

$$L(x, \lambda, \nu) = f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x)$$

- λ_i is the **Lagrange multiplier** associated with the ith inequality constraint $f_i(x) \leq 0$
- ν_i is the **Lagrange multiplier** associated with the ith equality constraint $h_i(x) = 0$
- The vectors λ and ν are called the **dual variables** or Lagrange multiplier vectors associated with the problem.

Lagrange dual function

The **Lagrange dual function** or just **dual function** $g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$ is the minimum value of the Lagrangian over x :

$$g(\lambda, \nu) = \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = \inf_{x \in \mathcal{D}} \left(f_0(x) + \sum_{i=1}^{m} \lambda_i f_i(x) + \sum_{i=1}^{p} \nu_i h_i(x) \right)$$

g is concave (even if the problem is not convex) and. When the Lagrangian is unbounded below in x, g can be $-\infty$.

Lower bound property

The dual function yields lower bounds on the optimal value of the problem.
If $\lambda \succeq 0$, then

$$g(\lambda, \nu) \leq p^\star$$

We have a lower bound that depends on some parameters λ, ν. We want to find the best lower bound that can be obtained from the Lagrange dual function.

Lagrange dual problem

The **Lagrange dual problem** associated with the standard form problem is

$$\begin{aligned} \text{maximize } & g(\lambda, \nu) \\ \text{s.t } & \lambda \succeq 0 \end{aligned}$$

- It is a a convex optimization problem. The optimal value is denoted $d^\star$.
- λ, ν are **dual feasible** if $\lambda \succeq 0$ and $g(\lambda, \nu) > -\infty$.

Weak and strong duality

Weak duality:

$$d^\star \leq p^\star$$

- It always holds (even if the original problem is not convex)
- It can be used to find nontrivial lower bounds for difficult problems.

Strong duality:

$$d^\star = p^\star$$

- It does not hold in general.
- It (usually) holds for convex problems bounds for difficult problems.
- The conditions that guarantee strong duality in convex problems are called **constraint qualifications**.

Slater's constraint qualification

Strong duality holds for a convex problem

$$\begin{aligned} \text{minimize } & f_0(x) \\ \text{s.t } & f_i(x) \leq 0, \ i = 1, \dots, m \\ & Ax = b \end{aligned}$$

if it is **strictly feasible**, meaning

$$\text{There exists } x \in \textbf{int } \mathcal{D} \text{ such that } f_i(x) < 0, \ i = 1, \dots, m, \quad Ax = b$$

- It also guarantees that the dual optimum is attained (if $p^\star > -\infty$).
- There are refinements (linear inequalities do not need to hold with strict inequality).
- There are other types of constraints qualifications.

Karush-Kuhn-Tucker (KKT) conditions

We assume the functions f_i, g_i are differentiable (no convexity assumptions). The following four conditions are called **Karush-Kuhn-Tucker (KKT) conditions**:

- primal constraints: $f_i(x) < 0,\ i = 1, ..., m,\ h_i(x) < 0,\ i = 1, ..., p$
- dual constraints : $\lambda \succeq 0$
- complementary slackness: $\lambda_i f_i(x) = 0,\ i = 1, ..., m$
- Stationarity : $\nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + \sum_{i=1}^{p} \nu_i \nabla h_i(x) = 0$

Optimality conditions

Necessary optimality conditions

For any problem, assuming strong duality holds, optimal points x, λ, ν must satisfy the KKT conditions.

Sufficient optimality conditions

For a convex problem, if x, λ, ν must satisfy the KKT conditions, then they are optimal.

Sufficient and necessary optimality conditions

For a convex problem, assuming Slater's condition is satisfied, x is optimal if and only if there are λ, ν that, together with x, satisfy the KKT conditions.