Lagrangian

standard form problem (not necessarily convex)

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & f_i(x) \leq 0, \quad i = 1, \ldots, m \\ & h_i(x) = 0, \quad i = 1, \ldots, p \end{array}$$

variable $x \in \mathbf{R}^n$, domain $\mathcal{D}$, optimal value $p^\star$

Lagrangian: $L : \mathbf{R}^n \times \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$, with $\mathbf{dom}\, L = \mathcal{D} \times \mathbf{R}^m \times \mathbf{R}^p$,

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

- weighted sum of objective and constraint functions

- λ_i is Lagrange multiplier associated with $f_i(x) \leq 0$

- ν_i is Lagrange multiplier associated with $h_i(x) = 0$

Lagrange dual function

Lagrange dual function: $g : \mathbf{R}^m \times \mathbf{R}^p \to \mathbf{R}$,

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

g is concave, can be $-\infty$ for some λ, ν

lower bound property: if $\lambda \succeq 0$, then $g(\lambda, \nu) \le p^\star$

proof: if $\tilde{x}$ is feasible and $\lambda \succeq 0$, then

$$
f_0(\tilde{x}) \ge L(\tilde{x}, \lambda, \nu) \ge \inf_{x \in \mathcal{D}} L(x, \lambda, \nu) = g(\lambda, \nu)
$$

minimizing over all feasible $\tilde{x}$ gives $p^\star \ge g(\lambda, \nu)$

The dual problem

Lagrange dual problem

$$
\begin{array}{ll}
\text{maximize} & g(\lambda, \nu) \\
\text{subject to} & \lambda \succeq 0
\end{array}
$$

- finds best lower bound on $p^\star$, obtained from Lagrange dual function

- a convex optimization problem; optimal value denoted $d^\star$

- λ, ν are dual feasible if $\lambda \succeq 0$, $(\lambda, \nu) \in \mathbf{dom}\, g$

- often simplified by making implicit constraint $(\lambda, \nu) \in \mathbf{dom}\, g$ explicit

example: standard form LP and its dual (page 5–5)

$$
\begin{array}{ll}
\text{minimize} & c^T x \\
\text{subject to} & Ax = b \\
& x \succeq 0
\end{array}
\qquad\qquad
\begin{array}{ll}
\text{maximize} & -b^T \nu \\
\text{subject to} & A^T \nu + c \succeq 0
\end{array}
$$

Weak and strong duality

weak duality: $d^\star \le p^\star$

- always holds (for convex and nonconvex problems)

- can be used to find nontrivial lower bounds for difficult problems

  for example, solving the SDP

$$\begin{array}{ll} \text{maximize} & -\mathbf{1}^T \nu \\ \text{subject to} & W + \mathbf{diag}(\nu) \succeq 0 \end{array}$$

  gives a lower bound for the two-way partitioning problem on page 5–7

strong duality: $d^\star = p^\star$

- does not hold in general

- (usually) holds for convex problems

- conditions that guarantee strong duality in convex problems are called **constraint qualifications**

Karush-Kuhn-Tucker (KKT) conditions

the following four conditions are called KKT conditions (for a problem with differentiable f_i, h_i):

1. primal constraints: $f_i(x) \leq 0$, $i = 1, \ldots, m$, $h_i(x) = 0$, $i = 1, \ldots, p$

2. dual constraints: $\lambda \succeq 0$

3. complementary slackness: $\lambda_i f_i(x) = 0$, $i = 1, \ldots, m$

4. gradient of Lagrangian with respect to x vanishes:

$$\nabla f_0(x) + \sum_{i=1}^{m} \lambda_i \nabla f_i(x) + \sum_{i=1}^{p} \nu_i \nabla h_i(x) = 0$$

from page 5–17: if strong duality holds and x, λ, ν are optimal, then they must satisfy the KKT conditions

Sufficient conditions for KKT at optima:

- **Slater's condition**: for **convex** problems, strong duality holds and KKT holds at optimum if there is a *strictly feasible* point x where all $f_i(x) < 0$ (or = 0 if f_i is affine).

- **LICQ** (linearly independent constraint qualification): even for **nonconvex** problems, KKT must hold at any local optima if the $\{\nabla f_i\}$ are linearly independent for *active* constraints
  - **Active** constraints: $f_i(x) = 0$