# 5. Duality

- Lagrange dual problem
- weak and strong duality
- geometric interpretation
- optimality conditions
- perturbation and sensitivity analysis
- examples
- generalized inequalities

# 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
- $\lambda_i$ is Lagrange multiplier associated with $f_i(x) \leq 0$
- $\nu_i$ 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 $\lambda$, $\nu$

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

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

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

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

# Least-norm solution of linear equations

$$\begin{array}{ll} \text{minimize} & x^T x \\ \text{subject to} & Ax = b \end{array}$$

**dual function**

- Lagrangian is $L(x, \nu) = x^T x + \nu^T (Ax - b)$
- to minimize $L$ over $x$, set gradient equal to zero:

$$\nabla_x L(x, \nu) = 2x + A^T \nu = 0 \quad \Longrightarrow \quad x = -(1/2)A^T \nu$$

- plug in in $L$ to obtain $g$:

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

a concave function of $\nu$

**lower bound property**: $p^\star \geq -(1/4)\nu^T A A^T \nu - b^T \nu$ for all $\nu$

# Standard form LP

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & Ax = b, \quad x \succeq 0 \end{array}$$

**dual function**

- Lagrangian is

$$\begin{aligned} L(x, \lambda, \nu) &= c^T x + \nu^T (Ax - b) - \lambda^T x \\ &= -b^T \nu + (c + A^T \nu - \lambda)^T x \end{aligned}$$

- $L$ is affine in $x$, hence

$$g(\lambda, \nu) = \inf_x L(x, \lambda, \nu) = \begin{cases} -b^T \nu & A^T \nu - \lambda + c = 0 \\ -\infty & \text{otherwise} \end{cases}$$

$g$ is linear on affine domain $\{(\lambda, \nu) \mid A^T \nu - \lambda + c = 0\}$, hence concave

**lower bound property**: $p^\star \geq -b^T \nu$ if $A^T \nu + c \succeq 0$

# Equality constrained norm minimization

$$\begin{array}{ll} \text{minimize} & \|x\| \\ \text{subject to} & Ax = b \end{array}$$

**dual function**

$$g(\nu) = \inf_x (\|x\| - \nu^T Ax + b^T \nu) = \begin{cases} b^T \nu & \|A^T \nu\|_* \leq 1 \\ -\infty & \text{otherwise} \end{cases}$$

where $\|v\|_* = \sup_{\|u\| \leq 1} u^T v$ is dual norm of $\| \cdot \|$

proof: follows from $\inf_x(\|x\| - y^T x) = 0$ if $\|y\|_* \leq 1$, $-\infty$ otherwise

- if $\|y\|_* \leq 1$, then $\|x\| - y^T x \geq 0$ for all $x$, with equality if $x = 0$
- if $\|y\|_* > 1$, choose $x = tu$ where $\|u\| \leq 1$, $u^T y = \|y\|_* > 1$:

$$\|x\| - y^T x = t(\|u\| - \|y\|_*) \to -\infty \quad \text{as } t \to \infty$$

**lower bound property:** $p^\star \geq b^T \nu$ if $\|A^T \nu\|_* \leq 1$

# Two-way partitioning

$$\begin{array}{ll} \text{minimize} & x^T W x \\ \text{subject to} & x_i^2 = 1, \quad i = 1, \ldots, n \end{array}$$

- a nonconvex problem; feasible set contains $2^n$ discrete points
- interpretation: partition $\{1, \ldots, n\}$ in two sets; $W_{ij}$ is cost of assigning $i$, $j$ to the same set; $-W_{ij}$ is cost of assigning to different sets

**dual function**

$$\begin{array}{rcl} g(\nu) = \inf\limits_x (x^T W x + \sum\limits_i \nu_i (x_i^2 - 1)) & = & \inf\limits_x x^T (W + \mathbf{diag}(\nu)) x - \mathbf{1}^T \nu \\ & = & \begin{cases} -\mathbf{1}^T \nu & W + \mathbf{diag}(\nu) \succeq 0 \\ -\infty & \text{otherwise} \end{cases} \end{array}$$

**lower bound property**: $p^\star \geq -\mathbf{1}^T \nu$ if $W + \mathbf{diag}(\nu) \succeq 0$

example: $\nu = -\lambda_{\min}(W)\mathbf{1}$ gives bound $p^\star \geq n\lambda_{\min}(W)$

# Lagrange dual and conjugate function

$$\begin{array}{ll} \text{minimize} & f_0(x) \\ \text{subject to} & Ax \preceq b, \quad Cx = d \end{array}$$

**dual function**

$$\begin{aligned} g(\lambda,\nu) &= \inf_{x\in\mathbf{dom}\, f_0} \left(f_0(x) + (A^T\lambda + C^T\nu)^T x - b^T\lambda - d^T\nu\right) \\ &= -f_0^*(-A^T\lambda - C^T\nu) - b^T\lambda - d^T\nu \end{aligned}$$

- recall definition of conjugate $f^*(y) = \sup_{x\in\mathbf{dom}\, f}(y^T x - f(x))$
- simplifies derivation of dual if conjugate of $f_0$ is kown

**example: entropy maximization**

$$f_0(x) = \sum_{i=1}^n x_i \log x_i, \qquad f_0^*(y) = \sum_{i=1}^n e^{y_i - 1}$$