About convex optimization with constraints

Alexandre Allauzen – Michèle Sebag
Thomas Schmitt
LIMSI – LRI

Oct. 7th, 2015

Reminder: optimisation with constraints

General case

$$\begin{cases} \text{Minimize} & f_0(\mathbf{z}) & \mathbf{z} \in \text{ domain } \mathcal{D} \\ \text{s.t.} & f_i(\mathbf{z}) \leq 0 & i = 1 \ldots m \\ & h_j(\mathbf{z}) = 0 & j = 1 \ldots p \end{cases} \quad (1)$$

Notation: $\mathbf{z}^*$ the solution of Pb (**??**)

$p^* = f_0(\mathbf{z}^*)$

Define the Lagrangian as:

$$L(\mathbf{z}, \lambda, \nu) = f_0(\mathbf{z}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{z}) + \sum_{j=1}^{p} \nu_j h_j(\mathbf{z})$$

with $\lambda_i \geq 0$

Reminder, 2

Define the Lagrange dual function as:

$$g(\lambda, \nu) = inf\{L(\mathbf{z}, \lambda, \nu), \mathbf{z} \in \mathcal{D}, \lambda \geq 0\} \tag{2}$$

Notation: $\lambda^*, \nu^* = \text{argmin } g(\lambda, \nu)$
$d^* = g(\lambda^*, \nu^*)$

Intuition

- λ_i, ν_j are the penalties paid for violating constraints
- For $\mathbf{z}$ feasible (i.e. $f_i(\mathbf{z}) \leq 0$ and $h_j(\mathbf{z}) = 0$),

$$L(\mathbf{z}, \lambda, \nu) \leq f_0(\mathbf{z})$$

- Hence $g(\lambda, \nu)$ provides a lower bound on the optimum p^*.
- L is linear in λ and ν: its minimum is very easy to compute;
- g is concave in λ and ν even if f_0, f_i and h_j are not convex in $\mathbf{z}$.

Reminder, 3

With

$$g(\lambda, \nu) = inf \left\{ f_0(\mathbf{z}) + \sum_{i=1}^{m} \lambda_i f_i(\mathbf{z}) + \sum_{j=1}^{p} \nu_j h_j(\mathbf{z}), \mathbf{z} \in \mathcal{D}, \lambda \geq 0 \right\}$$

Then

- Weak duality (always true)

$$d^* \leq p^*$$

- Strong duality: when $d^* = p^*$.
  usually holds for convex functions f_0 and f_i.

Reminder, 4

Constraint qualifications

= conditions guaranteeing strong duality in convex pbs.

Slater's conditions

if there exists $\mathbf{z}$ s.t. $f_i(\mathbf{z}) < 0$, then strong duality

Remark: if strong duality then

$$\begin{aligned} f_0(\mathbf{z}^*) &= inf_x \left\{ f_0(\mathbf{z}) + \sum_{i=1}^m \lambda_i^* f_i(\mathbf{z}) + \sum_{j=1}^p \nu_j^* h_j(\mathbf{z}) \right\} \\ &\leq f_0(\mathbf{z}^*) + \sum_{i=1}^m \lambda_i^* f_i(\mathbf{z}^*) + \sum_{j=1}^p \nu_j^* h_j(\mathbf{z}^*) \\ &\leq f_0(\mathbf{z}^*) \end{aligned}$$

Therefore for each $i = 1 \ldots m$, $\lambda_i^* f_i(\mathbf{z}^*) = 0$:

if $\lambda_i^* > 0$, $f_i(\mathbf{z}^*) = 0$ (saturated or tight constraint).

if $f_i(\mathbf{z}^*) < 0$ (loose constraint), $\lambda_i^* = 0$.

Reminder, 5

Karun-Kush-Tucker conditions on $\tilde{\mathbf{z}}, \tilde{\lambda}, \tilde{\nu}$

- Stationarity condition: f_i, h_j differentiable

$$\frac{\partial L(\mathbf{z}, \lambda, \nu)}{\partial \mathbf{z}}|_{\mathbf{z}=\tilde{\mathbf{z}}} = 0$$

- Slackness condition:

$$\tilde{\lambda}_i f_i(\tilde{\mathbf{z}}) = 0 \text{ for } i = 1 \ldots m$$

- Feasibility conditions

$$f_i(\tilde{\mathbf{z}}) \leq 0; \quad \tilde{\lambda}_i \geq 0$$

Theorem

$\Rightarrow$ If $\mathbf{z}^*$ is optimum of Pb (**??**), and λ^*, ν^* minimize g, then KKT conditions are satisfied for $(\mathbf{z}^*, \lambda^*, \nu^*)$.
$\Leftarrow$ If f_i convex and h_j affine, then if KKT conditions are satisfied for $\tilde{\mathbf{z}}, \tilde{\lambda}, \tilde{\nu}$, then $\tilde{\mathbf{z}}$ is optimum for Pb (**??**) and $\tilde{\lambda}, \tilde{\nu}$ minimize g.