About convex optimization with constraints

Alexandre Allauzen – Michèle Sebag
Thomas Schmitt
LIMSI – LRI

Oct. 7th, 2015
Reminder: optimisation with constraints

General case

\[
\begin{align*}
\text{Minimize} & \quad f_0(z) \quad z \in \text{domain } D \\
\text{s.t.} & \quad f_i(z) \leq 0 \quad i = 1 \ldots m \\
& \quad h_j(z) = 0 \quad j = 1 \ldots p
\end{align*}
\]

(1)

Notation: \( z^* \) the solution of Pb (??)
\[
p^* = f_0(z^*)
\]

Define the Lagrangian as:

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

with \( \lambda_i \geq 0 \)
Define the Lagrange dual function as:

\[ g(\lambda, \nu) = \inf \{ L(z, \lambda, \nu), z \in D, \lambda \geq 0 \} \]  

(2)

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

\( d^* = g(\lambda^*, \nu^*) \)

Intuition

- \( \lambda_i, \nu_j \) are the penalties paid for violating constraints
- For \( z \) feasible (i.e. \( f_i(z) \leq 0 \) and \( h_j(z) = 0 \)),
  \[ L(z, \lambda, \nu) \leq f_0(z) \]
- Hence \( g(\lambda, \nu) \) provides a lower bound on the optimum \( p^* \).
- \( L \) is linear in \( \lambda \) and \( \nu \): its minimum is very easy to compute;
- \( g \) is concave in \( \lambda \) and \( \nu \) even if \( f_0, f_i \) and \( h_j \) are not convex in \( z \).
Reminder, 3

With

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

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 $z$ s.t. $f_i(z) < 0$, then strong duality

**Remark:** if strong duality then

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

Therefore for each $i = 1 \ldots m$, $\lambda_i^* f_i(z^*) = 0$
  
  if $\lambda_i^* > 0$, $f_i(z^*) = 0$ (saturated or tight constraint).
  
  if $f_i(z^*) < 0$ (loose constraint), $\lambda_i^* = 0$. 

Reminder, 5

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

- Stationarity condition: $f_i, h_j$ differentiable

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

- Slackness condition:

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

- Feasibility conditions

\[
f_i(\tilde{z}) \leq 0; \quad \tilde{\lambda}_i \geq 0
\]

**Theorem**

⇒ If $z^*$ is optimum of Pb (??), and $\lambda^*, \nu^*$ minimize $g$, then KKT conditions are satisfied for $(z^*, \lambda^*, \nu^*)$.

⇐ If $f_i$ convex and $h_j$ affine, then if KKT conditions are satisfied for $\tilde{z}, \tilde{\lambda}, \tilde{\nu}$, then $\tilde{z}$ is optimum for Pb (??) and $\tilde{\lambda}, \tilde{\nu}$ minimize $g$. 