ECE 573 – Power System Operations and Control

4. Duality Theory of Nonlinear Programming

George Gross

Department of Electrical and Computer Engineering

University of Illinois at Urbana-Champaign
Consider a general nonlinear program

\[
\min f(x) \quad f(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}
\]

\[
\text{s.t.} \quad \ell(x) \leq 0 \quad \ell(\cdot) : \mathbb{R}^n \rightarrow \mathbb{R}^m
\]

\[
x \in \mathcal{I} \subset \mathbb{R}^n
\]

The Lagrangian function is defined for \( \lambda \in \mathbb{R}^m \)

\[
\mathcal{L}(x, \lambda) = f(x) + \lambda^T \ell(x)
\]


KUHN – TUCKER \((K-T)\) CONDITIONS

- The Kuhn-Tucker conditions are the necessary conditions of optimality, i.e., conditions implied by the existence of an optimal solution of the general nonlinear problem \((P)\)

- Consider \(\mathcal{I} \subset \mathbb{R}^n\) and assume \(f(\cdot)\) and \(\ell(\cdot)\) are differentiable and \((P)\) satisfies a constraint qualification \((CQ)\)
Two important cases of interest are:

(i) $\ell(\cdot)$ is linear; or

(ii) $\ell(\cdot)$ is convex and $\exists \ x \in S \ \exists \ l(x) < 0$

Definition: A function $\ell(\cdot)$ is convex $\iff$

$\forall x, y \in S$ and $\forall \alpha \in [0,1]$,

$\ell[\alpha x + (1 - \alpha) y] \leq \alpha \ell(x) + (1 - \alpha) \ell(y)$
$K-T$ CONDITONS

- $K-T$ conditions are the necessary conditions of optimality

- A feasible point $x^*$ of $(P)$ is optimal, then there exists $\lambda^* \geq 0$ that satisfies the stationarity and complementary slackness conditions:

$$0 \leq A x^* - b$$

$$c(x^*) = 0$$

$$0 \leq A^T \lambda^*$$
**K-T CONDITIONS**

(i) **stationarity**

\[ \nabla_x \mathcal{L}(x, \lambda) \bigg|_{x^*, \lambda^*} = 0 \]

(ii) **complementary slackness (c-s)**

\[
\begin{pmatrix} \lambda^* \end{pmatrix}^T \ell(x^*) = 0 \iff \lambda^*_i \ell_i(x^*) = 0 \quad \forall i
\]

with

\[ \lambda^* \geq 0 \iff \lambda^*_i \geq 0 \]

There are a number of cases for which the constraint qualification guarantees that the **K-T theorem** holds
THE DUAL FUNCTION

The dual function of \((P)\) is defined to be

\[
h(\lambda) \triangleq \min_{x \in \mathcal{I}} \mathcal{L}(x, \lambda)
\]

Let

\[
\mathcal{D} \triangleq \{ \lambda : h(\lambda) \text{ exists and } \lambda \geq 0 \}
\]

Then, the dual problem is

\[
\begin{aligned}
\max & \quad h(\lambda) \\
\text{s.t.} & \quad \lambda \in \mathcal{D}
\end{aligned}
\]

\((D)\)
SADDLE POINT DEFINITION

A function \( \mathcal{L} (\cdot, \cdot) : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R} \) has a \textit{saddle point} at \( (x_0, \lambda_0) \) \iff there exists \( \varepsilon > 0 \) \suchthat \( \forall x \) with

\[
\| x - x_0 \| < \varepsilon \quad \text{and} \quad \forall \lambda \quad \text{with} \quad \| \lambda - \lambda_0 \| < \varepsilon
\]

\( \mathcal{L} (x_0, \lambda) \leq \mathcal{L} (x_0, \lambda_0) \leq \mathcal{L} (x, \lambda_0) \)

\( \text{maximum of } \mathcal{L} \text{ w.r.t. } \lambda \)

\( \text{minimum of } \mathcal{L} \text{ w.r.t. } x \)
SADDLE POINT : EXAMPLES

http://en.wikipedia.org/wiki/Saddle_point
SADDLE POINT: EXAMPLES
GLOBAL OPTIMALITY CONDITIONS

Let \( x^* \in \mathcal{X} \) and \( \lambda^* \in \mathcal{D} \). For the primal-dual problems \((P)\) and \((D)\), we have

(i) minimality

\[ x^* \text{ minimizes } \mathcal{L}(x, \lambda^*), x \in \mathcal{X} \]

\[ h(\lambda^*) = \min_{x \in \mathcal{X}} \mathcal{L}(x, \lambda^*) = \mathcal{L}(x^*, \lambda^*) \]

(ii) feasibility

\[ \ell(x^*) \leq 0 \]

(iii) complementary slackness \((c-s)\) conditions

\[ (\lambda^*)^T \ell(x^*) = 0 \]
GLOBAL OPTIMALITY IMPLICATIONS

☐ Minimality and the objective of \((D)\) imply

\[
\mathcal{L}(x^*, \lambda) \leq \mathcal{L}(x^*, \lambda^*) \leq \mathcal{L}(x, \lambda^*)
\]  

\((x^*, \lambda^*)\) is a saddle point of \(\mathcal{L}(x, \lambda)\)

☐ Minimality and complementary slackness imply

\(x^*\) is optimal for \((P)\)

☐ Complementary slackness and (*) imply

\(\lambda^*\) is optimal for \((D)\)
LEMMA

- Let $x \in \mathcal{X}$ be arbitrary and let $\overline{x}$ be feasible for $(P)$.

- Then, $\forall \lambda \in \mathcal{D}$

\[
h(\lambda) = \min_{y \in \mathcal{Y}} \mathcal{L}(y, \lambda) \leq \mathcal{L}(x, \lambda) = f(x) + \lambda^T \ell(x) \leq f(x)
\]

- The objective of the dual is bounded from above by the value of the objective of the primal at any arbitrary point.
STRONG DUALITY THEOREM

\( \mathbf{x}^* \) and \( \lambda^* \) satisfy the global optimality conditions \( \iff \)

(i) \( \mathbf{x}^* \) is feasible in \( (P) \)

(ii) \( \lambda^* \) is feasible in \( (D) \)

(iii) \( f(\mathbf{x}^*) = h(\lambda^*) = \max_{\lambda \in \mathcal{D}} h(\lambda) \)

\( \square \) Corollary: For \( f(\cdot) \) and \( h(\cdot) \) differentiable and convex functions, the global optimality conditions are equivalent to the \( K-T \) conditions
A SPECIAL PROBLEM

\[
\begin{align*}
    \min & \quad f(x) \\
    \text{s.t.} & \quad \ell(x) \leq 0 \\
    \quad & \quad x \geq 0
\end{align*}
\]

Define for \((P_s)\) the Lagrangian

\[
\tilde{\mathcal{L}}(x, \lambda, \mu) = f(x) + \lambda^T \ell(x) + \mu^T (-x) = \mathcal{L}(x, \lambda) - \mu^T x
\]
**K–T CONDITIONS FOR A SPECIAL PROBLEM**

There exist dual variables $\lambda^* \geq 0$, $\mu^* \geq 0$ by stationarity

\[ \nabla_x \tilde{\mathcal{L}} \bigg|_{x^*, \lambda^*, \mu^*} = \nabla_x f \bigg|_{x^*} + (\lambda^*)^T \nabla_x \ell \bigg|_{x^*} + (\mu^*)^T (-1) = 0^T \]

\[ \nabla_x \mathcal{L} \bigg|_{x^*, \lambda^*} = \nabla_x f \bigg|_{x^*} + (\lambda^*)^T \nabla_x \ell \bigg|_{x^*} = (\mu^*)^T \geq 0^T \]

\[ \frac{\partial}{\partial x_i} \mathcal{L} \bigg|_{x^*, \lambda} \geq 0 \quad \forall \ i \]
Furthermore, the $c-s$ conditions imply

\[ \lambda^* T \ell(x^*) - \mu^* T x^* = 0 \]

which may only hold if $\mu^* T x^* = 0$ and $\lambda^* T \ell(x^*) = 0$

\[ \nabla_x \tilde{\mathcal{L}} \bigg|_{x^*, \lambda^*, \mu^*} = \nabla_x \mathcal{L} \bigg|_{x^*, \lambda^*, - (\mu^*)^T} = 0^T \]

\[ \nabla_x \mathcal{L} \bigg|_{x^*, \lambda^*, x} = \mu^* T x^* = 0 \]

The $c-s$ conditions are

\[ \nabla_x \mathcal{L} \bigg|_{x^*, \lambda^*} x^* = 0 \quad \left(\lambda^* \right)^T \ell(x^*) = 0 \]