DUALITY IN OPTIMIZATION

Prof. Dan A. Simovici

UMB

Primal and Dual Problems

Let $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, $\mathbf{g} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ and $\mathbf{h} : \mathbb{R}^n \longrightarrow \mathbb{R}^\ell$ and let $\theta : \mathbb{R}^{m+\ell} \longrightarrow \mathbb{R}$ given by:

$$\theta(\mathbf{u}, \mathbf{v}) = \inf_{\mathbf{x} \in X} L(\mathbf{x}, \mathbf{u}, \mathbf{v}),$$

where

$$L(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \mathbf{u}'\mathbf{g}(\mathbf{x}) + \mathbf{v}'h(\mathbf{x})$$

is the Lagrangean function.

Primal Problem:

minimize $f(\mathbf{x})$

subject to: $\mathbf{g}(\mathbf{x}) \leqslant \mathbf{0}$

$\mathbf{h}(\mathbf{x}) = \mathbf{0}$

$\mathbf{x} \in X$

Dual Problem:

maximize $\theta(\mathbf{u}, \mathbf{v})$

subject to: $\mathbf{u} \geqslant \mathbf{0}$

Weak Duality Theorem

Theorem

Le $\mathbf{x} \in \mathbb{R}^n$ *be a feasible point of the primal problem, that is,*

$$\mathbf{x} \in X, \mathbf{g}(\mathbf{x}) \leqslant \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}.$$

Let $\begin{pmatrix} \mathbf{u} \\ \mathbf{v} \end{pmatrix}$ *be a feasible solution to the dual problem, that is,* $\mathbf{u} \geqslant \mathbf{0}$*. Then,*

$$\theta(\mathbf{u}, \mathbf{v}) \leqslant f(\mathbf{x}).$$

Proof: If $u \geqslant 0$, then $f(\mathbf{x}) + \mathbf{u}'\mathbf{g}(\mathbf{x}) + \mathbf{v}'\mathbf{h}(\mathbf{x}) \leqslant f(\mathbf{x})$ because $\mathbf{g}(\mathbf{x}) \leqslant \mathbf{0}$ and $\mathbf{h}(\mathbf{x}) = \mathbf{0}$. Therefore,

$$\theta(\mathbf{u}, \mathbf{v}) = \inf_{\mathbf{x}}(f(\mathbf{x}) + \mathbf{u}'\mathbf{g}(\mathbf{x}) + \mathbf{v}'\mathbf{h}(\mathbf{x})) \leqslant \inf_{\mathbf{x}} f(\mathbf{x}),$$

which implies $\theta(\mathbf{u}, \mathbf{v}) \leqslant f(\mathbf{x})$.

Corollaries

The weak duality theorem can be used to obtain lower bounds on optimal values of difficult problems.

A The dual optimum (max) is a lower bound of the primal optimum (min), that is,

$$\inf\{f(\mathbf{x}) \mid \mathbf{x} \in X, \mathbf{g}(\mathbf{x}) \leqslant \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\} \geqslant \sup\{\theta(\mathbf{u}, \mathbf{v}) \mid \mathbf{u} \geqslant \mathbf{0}\}.$$

B If $f(\overline{\mathbf{x}}) \leqslant \theta(\overline{\mathbf{u}}, \overline{\mathbf{v}})$, where $\overline{\mathbf{u}} \geqslant \mathbf{0}$ and $\overline{\mathbf{x}} \in \{\mathbf{x} \in X, \mathbf{g}(\mathbf{x}) \leqslant \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\}$, then $\overline{\mathbf{x}}$ and $\begin{pmatrix} \overline{\mathbf{u}} \\ \overline{\mathbf{v}} \end{pmatrix}$ solve the primal and dual problems, respectively.

C If $\inf\{f(\mathbf{x}) \mid \mathbf{x} \in X, \mathbf{g}(\mathbf{x}) \leqslant \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\} = -\infty$, then $\theta(\mathbf{u}, \mathbf{v}) = -\infty$.

Recall

Let S be a nonempty convex subset of $\mathbb{R}^n$ and $\overline{\mathbf{x}} \not\in S$. Then, there exists a $\mathbf{p} \neq \mathbf{0}$ such that $\mathbf{p}'(\mathbf{x} - \overline{\mathbf{x}}) \geqslant \mathbf{0}$ for every $\mathbf{x} \in \mathbf{K}(S)$.
In particular, if $\mathbf{0} \not\in S$, there exists $\mathbf{p} \neq \mathbf{0}$ such that $\mathbf{p}'\mathbf{x} \geqslant \mathbf{0}$ for $\mathbf{x} \in \mathbf{K}(S)$.

Theorem

Let X be nonempty and convex subset of $\mathbb{R}^n$. Let $\alpha : \mathbb{R}^n \longrightarrow \mathbb{R}$ and $\mathbf{g} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be convex and let $\mathbf{h} : \mathbb{R}^n \longrightarrow \mathbb{R}^\ell$ be given by $\mathbf{h}(\mathbf{x}) = A\mathbf{x} - \mathbf{b}$. If the system (S1)

$$\alpha(\mathbf{x}) < 0, \mathbf{g}(\mathbf{x}) \leqslant \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}$$

has no solution, then the system (S2):

$$\begin{gathered} u_0\alpha(\mathbf{x}) + \mathbf{u}'\mathbf{g}(\mathbf{x}) + \mathbf{v}'\mathbf{h}(\mathbf{x}) \geqslant \mathbf{0}, \\ \begin{pmatrix} u_0 \\ \mathbf{u} \end{pmatrix} \geqslant \mathbf{0}, \begin{pmatrix} u_0 \\ \mathbf{u} \\ \mathbf{v} \end{pmatrix} \neq \mathbf{0}, \end{gathered}$$

has a solution. The converse holds when $u_0 > 0$.

Proof

Suppose that (S1) has no solution and consider the set

$$L = \left\{ \begin{pmatrix} p \\ \mathbf{q} \\ \mathbf{r} \end{pmatrix} \middle| p > \alpha(\mathbf{x}), \mathbf{q} \geqslant g(\mathbf{x}), \mathbf{r} = \mathbf{h}(\mathbf{x}), \mathbf{x} \in X \right\}$$

The set L is convex. Since (S1) has no solution, $\begin{pmatrix} 0 \\ \mathbf{0} \\ \mathbf{0} \end{pmatrix} \not\in L$. By the result recalled earlier, there exists $\begin{pmatrix} u_0 \\ \mathbf{u} \\ \mathbf{v} \end{pmatrix}$ such that $u_0 p + \mathbf{u}'\mathbf{q} + \mathbf{v}'\mathbf{r} \geqslant \mathbf{0}$ for each $\begin{pmatrix} p \\ \mathbf{q} \\ \mathbf{r} \end{pmatrix} \in L$

Fix $\mathbf{x} \in X$. Since p and $\mathbf{q}$ can be made arbitrarily large, we have $u_0 p + \mathbf{u}'\mathbf{q} + \mathbf{v}'\mathbf{r} \geqslant \mathbf{0}$ only if $u_0 \geqslant 0$ and $\mathbf{u} \geqslant \mathbf{0}$. This shows that (S2) has a solution.

Conversely, suppose that (S2) has a solution $\begin{pmatrix} u_0 \\ \mathbf{u} \\ \mathbf{v} \end{pmatrix}$ such that

$$u_0 > 0 \text{ and } \mathbf{u} \geqslant 0$$

that satisfies

$$u_0\alpha(\mathbf{x}) + \mathbf{u}'\mathbf{g}(\mathbf{x}) + \mathbf{v}'\mathbf{h}(\mathbf{x}) \geqslant \mathbf{0},$$

for $\mathbf{x} \in X$.
Let $\mathbf{x}$ be such that $\mathbf{g}(\mathbf{x}) \leqslant \mathbf{0}$ and $\mathbf{h}(\mathbf{x}) = \mathbf{0}$. Since $\mathbf{u} \geqslant \mathbf{0}$, it follows that $u_0\alpha(\mathbf{x}) \geqslant 0$. Since $u_0 > 0$, it follows that $\alpha(\mathbf{x}) \geqslant 0$, so (S1) has no solution.

Framework for Strong Duality

Let $X \subseteq \mathbb{R}^n$ be a nonempty convex subset.

- $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, $\mathbf{g} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ be convex functions and let $\mathbf{h} : \mathbb{R}^n \longrightarrow \mathbb{R}$ given by $\mathbf{h}(\mathbf{x}) = A\mathbf{x} - \mathbf{b}$.
- there exists $\hat{\mathbf{x}}$ such that $\mathbf{g}(\hat{\mathbf{x}}) \leqslant \mathbf{0}$ and $\mathbf{h}(\hat{\mathbf{x}}) = \mathbf{0}$;
- $\mathbf{0} \in \mathbf{h}(X) = \{\mathbf{h}(\mathbf{x}) \mid \mathbf{x} \in X\}$.

Strong Duality:

$$\sup\{\theta(\mathbf{u}, \mathbf{v}) \mid \mathbf{u} \geqslant \mathbf{0}\} = \inf\{f(\mathbf{x}) \mid \mathbf{x} \in X, \mathbf{g}(\mathbf{x}) \leqslant \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\}$$

Strong Duality Theorem

Theorem

Let X *be a nonempty convex subset of* $\mathbb{R}^n$, $f : \mathbb{R}^n \longrightarrow \mathbb{R}$, *and* $\mathbf{g} : \mathbb{R}^n \longrightarrow \mathbb{R}^m$ *be a convex function and let* $\mathbf{h} : \mathbb{R}^n \longrightarrow \mathbb{R}$ *be a function given by* $\mathbf{h}(\mathbf{x}) = A\mathbf{x} - \mathbf{b}$.

Suppose that the following constraint qualification holds:

- *there exists* $\hat{\mathbf{x}}$ *such that* $\mathbf{g}(\hat{\mathbf{x}}) \leqslant \mathbf{0}$ *and* $\mathbf{h}(\hat{\mathbf{x}}) = \mathbf{0}$;
- $\mathbf{0} \in \mathbf{I}(\mathbf{h}(X))$, *where* $\mathbf{h}(X) = \{\mathbf{h}(\mathbf{x}) \mid \mathbf{x} \in X\}$.

Then,

$$\inf\{f(\mathbf{x}) \mid \mathbf{x} \in X, \mathbf{g}(\mathbf{x}) \leqslant \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\} = \sup\{\theta(\mathbf{u}, \mathbf{v}) \mid \mathbf{u} \geqslant \mathbf{0}\}$$

Proof

Let

$$\gamma = \inf\{f(\mathbf{x}) \mid \mathbf{x} \in X, \mathbf{g}(\mathbf{x}) \leqslant \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\}.$$

If $\gamma = -\infty$, then by Corollary (C), $\sup\{\theta(\mathbf{u}, \mathbf{v}) \mid \mathbf{u} \geqslant \mathbf{0}\} = -\infty$ and the equality holds.
Suppose γ is finite and consider the system:

$$f(\mathbf{x}) - \gamma < 0, \mathbf{g}(\mathbf{x}) \leqslant \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}, \mathbf{x} \in X,$$

which has so solution. Therefore, by the Theorem on slide 7, there exists $\begin{pmatrix} u_0 \\ \mathbf{u} \\ \mathbf{v} \end{pmatrix}$ such that

$$u_0(f(\mathbf{x}) - \gamma) + \mathbf{u}'\mathbf{g}(\mathbf{x}) + \mathbf{v}'\mathbf{h}(\mathbf{x}) \geqslant 0, \tag{1}$$

for $\mathbf{x} \in X$.

Proof (cont'd)

We claim that $u_0 > 0$. Suppose that this is not the case, that is, $u_0 = 0$. By the constraint qualifications there exists $\hat{\mathbf{x}} \in X$ such that $\mathbf{g}(\hat{\mathbf{x}}) < 0$ and $\mathbf{h}(\hat{\mathbf{x}}) = \mathbf{0}$. The inequality $u_0(f(\hat{\mathbf{x}}) - \gamma) + \mathbf{u}'\mathbf{g}(\hat{\mathbf{x}}) + \mathbf{v}'\mathbf{h}(\hat{\mathbf{x}}) \geqslant 0$, reduces to $\mathbf{u}'\mathbf{g}(\hat{\mathbf{x}}) \geqslant 0$. Since $\mathbf{g}(\hat{\mathbf{x}}) < 0$ and $\mathbf{u} \geqslant \mathbf{0}$, $\mathbf{u}'\mathbf{g}(\hat{\mathbf{x}}) \geqslant 0$ is possible only if $\mathbf{u} = \mathbf{0}$.

Again, from $u_0(f(\mathbf{x}) - \gamma) + \mathbf{u}'\mathbf{g}(\mathbf{x}) + \mathbf{v}'\mathbf{h}(\mathbf{x}) \geqslant 0$, $u_0 = 0$ and $\hat{\mathbf{u}} = \mathbf{0}$ implies $\mathbf{v}'\mathbf{h}(\mathbf{x}) \geqslant 0$ for $\mathbf{x} \in X$.

Since $\mathbf{0} \in \mathbf{I}(\mathbf{h}(X))$, we can move from $\mathbf{0}$ in the direction $-\lambda\mathbf{v}$ and still remain in $\mathbf{h}(X)$, that is, there exists $\mathbf{x} \in X$ such that $\mathbf{h}(\mathbf{x}) = -\lambda\mathbf{v}$, where $\lambda > 0$. Therefore,

$$0 \leqslant \mathbf{v}'\mathbf{h}(\mathbf{x}) = -\lambda \parallel \mathbf{v} \parallel^2,$$

which implies $\mathbf{v} = \mathbf{0}$. We have shown that $u_0 = 0$ implies $\begin{pmatrix} u_0 \\ \mathbf{u} \\ \mathbf{v} \end{pmatrix} = \mathbf{0}$,

which is impossible. Hence $u_0 > 0$.

Proof (cont'd)

Dividing Inequality (1) by u_0 we get

$$f(\mathbf{x}) + \overline{\mathbf{u}}'\mathbf{g}(\mathbf{x}) + \overline{\mathbf{v}}'\mathbf{h}(\mathbf{x}) \geqslant \gamma \tag{2}$$

for $\mathbf{x} \in X$, where $\overline{\mathbf{u}} = \frac{1}{u_0}\mathbf{u}$ and $\overline{\mathbf{v}} = \frac{1}{u_0}\mathbf{v}$. This shows that

$$\theta(\overline{\mathbf{u}}, \overline{\mathbf{v}}) = \inf_{\mathbf{x} \in X} f(\mathbf{x}) + \overline{\mathbf{u}}'\mathbf{g}(\mathbf{x}) + \overline{\mathbf{v}}'\mathbf{h}(\mathbf{x}) \geqslant \gamma.$$

Therefore, $\theta(\overline{\mathbf{u}}, \overline{\mathbf{v}}) = \gamma$, so $\overline{\mathbf{u}}$ and $\overline{\mathbf{v}}$ solve the dual problem.

Proof (cont'd)

Suppose that $\overline{\mathbf{x}}$ is an optimal solution to the primal problem, that is, $\overline{\mathbf{x}} \in X$, $\mathbf{g}(\overline{\mathbf{x}}) \leqslant \mathbf{0}$, $\mathbf{h}(\overline{\mathbf{x}}) = \mathbf{0}$ and $f(\overline{\mathbf{x}}) = \gamma$. Choosing $\mathbf{x} = \overline{\mathbf{x}}$ in Inequality (2):

$$f(\mathbf{x}) + \overline{\mathbf{u}}'\mathbf{g}(\mathbf{x}) + \overline{\mathbf{v}}'\mathbf{h}(\mathbf{x}) \geqslant \gamma$$

we get $\overline{\mathbf{u}}'g(\overline{\mathbf{x}}) \geqslant 0$.
Since $\overline{\mathbf{u}} \geqslant \mathbf{0}$ and $\mathbf{g}(\overline{\mathbf{x}}) \leqslant \mathbf{0}$, we have $\overline{\mathbf{u}}'\mathbf{g}(\overline{\mathbf{x}}) = 0$.

Example (non-convex X)

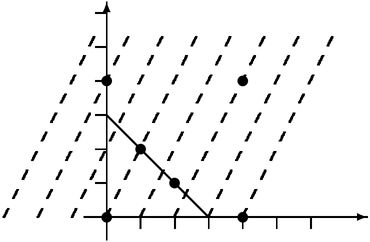
Consider the primal problem:

- minimize $-2x_1 + x_2$;
- subject to $x_1 + x_2 - 3 = 0$, $\mathbf{x} \in X$,

where

$$X = \left\{ \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \begin{pmatrix} 0 \\ 4 \end{pmatrix}, \begin{pmatrix} 4 \\ 0 \end{pmatrix}, \begin{pmatrix} 4 \\ 4 \end{pmatrix}, \begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 2 \\ 1 \end{pmatrix} \right\}.$$

The optimal solution to the primal is $\begin{pmatrix} 2 \\ 1 \end{pmatrix}$ with objective value -3.



$(2, 1)$ is the optimal solution of the primal; the dual objective function is

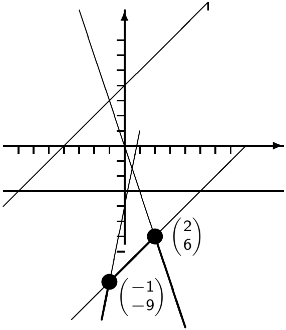
$$\theta(v) = \min\{(-2x_1 + x_2) + v(x_1 + x_2 - 3) \mid \mathbf{x} \in X\}.$$

Example (cont'd)

$$\begin{array}{cccccc} \begin{pmatrix}0\\0\end{pmatrix} & \begin{pmatrix}0\\4\end{pmatrix} & \begin{pmatrix}4\\4\end{pmatrix} & \begin{pmatrix}4\\0\end{pmatrix} & \begin{pmatrix}1\\2\end{pmatrix} & \begin{pmatrix}2\\1\end{pmatrix} \\ -3v & 4+v & -4+5v & -8+v & 0 & -3 \end{array}$$

This implies

$$\theta(v) = \begin{cases} -4+5v & \text{if } v \leqslant -1 \\ -8+v & \text{if } -1 \leqslant v \leqslant 2 \\ -3v & \text{if } v \geqslant 2 \end{cases}$$



Example (non-convex X)

Let $A \in \mathbb{R}^{n \times n}$ be a matrix. Consider the problem

- minimize $\mathbf{x}' A \mathbf{x}$;
- subject to $x_i^2 = 1$, that is, $x_i \in \{-1, 1\}$ for $1 \leqslant i \leqslant n$.

The feasible set contains 2^n points.
The Lagrangean function is

$$\begin{aligned} L(\mathbf{x}, \mathbf{v}) &= \mathbf{x}' A \mathbf{x} + \sum_{j=1}^{n} v_j (x_j^2 - 1) \\ &= \mathbf{x}'(A + \mathrm{diag}(v_1, \ldots, v_n))\mathbf{x} - \mathbf{1}'\mathbf{v}. \end{aligned}$$

The dual function is

$$\theta(\mathbf{v}) = \inf_{\mathbf{x}} L(\mathbf{x}, \mathbf{v}) = \begin{cases} -\mathbf{1}'\mathbf{v} & \text{if } A + \mathrm{diag}(v_1, \ldots, v_n) \geqslant \mathbf{0} \\ -\infty & \text{otherwise}, \end{cases}$$

provides lower bounds on the optimal value of the primal problem. For example, we can take $\mathbf{v} = -\lambda_{min}\mathbf{1}$, where λ_{min} is the smallest eigenvalue of A. This is dually feasible because $A + \mathrm{diag}(\mathbf{v}) = A - \lambda_{min} I_n \geqslant \mathbf{0}$, so the optimal value of the primal p^* satisfies

$$p^* \geqslant -\mathbf{1}'\mathbf{v} = n\lambda_{min}.$$

Example

Consider the problem

- minimize $\parallel \mathbf{x} \parallel$ for $\mathbf{x} \in \mathbb{R}^n$;
- subjected to $A\mathbf{x} = \mathbf{b}$, where $A \in \mathbb{R}^{m \times n}$.

Minimizing $\parallel \mathbf{x} \parallel$ is tantamount to minimizing $\mathbf{x}'\mathbf{x}$ and this is the function we are using.
The Lagrangean is

$$L(\mathbf{x}, \mathbf{v}) = \mathbf{x}'\mathbf{x} + \mathbf{v}'(A\mathbf{x} - \mathbf{b})$$

and is a quadratic convex function in $\mathbf{x}$. The objective function of the dual is $\theta(\mathbf{v}) = \inf_{\mathbf{x}} L(\mathbf{x}, \mathbf{v})$, where $\mathbf{v} \in \mathbb{R}^m$.
The minimum of L is determined from

$$\nabla_{\mathbf{x}} L(\mathbf{x}, \mathbf{v}) = 2\mathbf{x} + \mathbf{v}'A = \mathbf{0},$$

so $\mathbf{x} = -\frac{1}{2}\mathbf{v}'A$. Therefore,

$$\theta(\mathbf{v}) = -\frac{1}{4}\mathbf{v}'AA'\mathbf{v} - \mathbf{b}'\mathbf{v},$$

which is a concave quadratic function with domain $\mathbb{R}^m$. The weak duality theorem states that

$$-\frac{1}{4}\mathbf{v}'AA'\mathbf{v} - \mathbf{b}'\mathbf{v} \leqslant \inf\{\mathbf{x}'\mathbf{x} \mid A\mathbf{x} = \mathbf{b}\}.$$

Saddle Point Theorem

Theorem

Let $X \neq \emptyset$ *be a subset of* $\mathbb{R}^n$ *and let* $f : \mathbb{R}^n \longrightarrow \mathbb{R}, g : \mathbb{R}^n \longrightarrow \mathbb{R}, h : \mathbb{R}^n \longrightarrow \mathbb{R}$. *Suppose that there exists* $\overline{\mathbf{x}} \in X$ *and* $\overline{\mathbf{u}}, \overline{\mathbf{v}}$ *with* $\overline{\mathbf{u}} \geqslant \mathbf{0}$ *such that*

$$\Phi(\overline{\mathbf{x}}, \mathbf{u}, \mathbf{v}) \leqslant \Phi(\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\mathbf{v}}) \leqslant \Phi(\mathbf{x}, \overline{\mathbf{u}}, \overline{\mathbf{v}}) \tag{3}$$

for all $\mathbf{x} \in X$ *and all* $\mathbf{u}, \mathbf{v}$ *with* $\mathbf{u} \geqslant \mathbf{0}$, *where*

$$\Phi(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \mathbf{u}'\mathbf{g}(\mathbf{x}) + \mathbf{v}'h(\mathbf{x}).$$

Then $\overline{\mathbf{x}}$ *solves the primal problem, and* $\mathbf{u}, \mathbf{v}$ *solve the dual problem. Conversely, suppose that* X, f, g *are convex,* $\mathbf{h}(\mathbf{x}) = A\mathbf{x} - \mathbf{b}$, $\mathbf{0} \in \mathbf{I}(\mathbf{h}(X))$, *there exists* $\hat{\mathbf{x}} \in X$ *such that* $\mathbf{g}(\hat{\mathbf{x}}) < \mathbf{0}$ *and* $\mathbf{h}(\hat{\mathbf{x}}) = \mathbf{0}$. *If* $\overline{\mathbf{x}}$ *is an optimal solution to the primal problem, then there exist* $\overline{\mathbf{u}}, \overline{\mathbf{v}}$ *with* $\overline{\mathbf{u}} \geqslant \mathbf{0}$ *so that the inequalities (3) hold.*

Suppose that there exists $\overline{\mathbf{x}} \in X$ and $\overline{\mathbf{u}}, \overline{\mathbf{v}}$ with $\overline{\mathbf{u}} \geqslant \mathbf{0}$ such that

$$\Phi(\overline{\mathbf{x}}, \mathbf{u}, \mathbf{v}) \leqslant \Phi(\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\mathbf{v}}) \leqslant \Phi(\mathbf{x}, \overline{\mathbf{u}}, \overline{\mathbf{v}}).$$

Since

$$f(\overline{\mathbf{x}}) + \mathbf{u}'\mathbf{g}(\overline{\mathbf{x}}) + \mathbf{v}'\mathbf{h}(\overline{\mathbf{x}}) = \Phi(\overline{\mathbf{x}}, \mathbf{u}, \mathbf{v}) \leqslant \Phi(\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\mathbf{v}}) = f(\overline{\mathbf{x}}) + \overline{\mathbf{u}}'\mathbf{g}(\overline{\mathbf{x}}) + \overline{\mathbf{v}}'\mathbf{h}(\overline{\mathbf{x}}),$$

for all $\mathbf{u} \geqslant \mathbf{0}$ and all $\mathbf{v} \in \mathbb{R}^n$ it follows that

$$(\mathbf{u} - \overline{\mathbf{u}})'\mathbf{g}(\overline{\mathbf{x}}) \leqslant (\overline{\mathbf{v}} - \mathbf{v})'\mathbf{h}(\overline{\mathbf{x}}),$$

for all $\mathbf{u} \geqslant \mathbf{0}$ and all $\mathbf{v} \in \mathbb{R}^n$, which implies $\mathbf{g}(\overline{\mathbf{v}}) \leqslant \mathbf{0}$ and $\mathbf{h}(\overline{\mathbf{x}}) = \mathbf{0}$. Thus, $\overline{\mathbf{x}}$ is a feasible solution of P.
By taking $\mathbf{u} = \mathbf{0}$, it follows that $\overline{\mathbf{u}}'\mathbf{g}(\overline{\mathbf{x}}) \geq \mathbf{0}$. Since $\overline{\mathbf{u}} \geqslant \mathbf{0}$ and $\mathbf{g}(\overline{\mathbf{x}}) \leqslant \mathbf{0}$, we have $\overline{\mathbf{u}}'\mathbf{g}(\overline{\mathbf{x}}) = \mathbf{0}$.

By the inequalities of the hypothesis we get

$$\begin{aligned} f(\overline{\mathbf{x}}) &= f(\overline{\mathbf{x}}) + \overline{\mathbf{u}}' g(\overline{\mathbf{x}}) + \overline{\mathbf{v}}' \mathbf{h}(\overline{\mathbf{x}}) = \Phi(\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\mathbf{v}}) \\ &\leqslant \Phi(\mathbf{x}, \overline{\mathbf{u}}, \overline{\mathbf{v}}) = f(\mathbf{x}) + \overline{\mathbf{u}}' g(\mathbf{x}) + \overline{\mathbf{v}}' \mathbf{h}(\mathbf{x}) \end{aligned}$$

for every $\mathbf{x} \in X$. This implies $f(\overline{\mathbf{x}}) \leqslant \theta(\mathbf{u}, \mathbf{v})$. Since $\overline{\mathbf{x}}$ is feasible for the primal problem and $\overline{\mathbf{u}} \geqslant \mathbf{0}$, it follow that $\overline{\mathbf{x}}$ and $\overline{\mathbf{u}}, \overline{\mathbf{v}}$ are optimal for the primal and the dual problem, respectively.

Conversely, suppose that $\overline{\mathbf{x}}$ is optimal for the primal problem. By the Strong Duality Theorem, there exist $\overline{\mathbf{u}}, \overline{\mathbf{v}}$ with $\overline{\mathbf{u}} \leqslant \mathbf{0}$ such that $f(\overline{\mathbf{x}}) = \theta(\overline{\mathbf{u}}, \overline{\mathbf{v}})$ and $\overline{\mathbf{u}}'\mathbf{g}(\overline{\mathbf{x}}) = \mathbf{0}$. so

$$f(\overline{\mathbf{x}}) = \theta(\overline{\mathbf{u}}, \overline{\mathbf{v}}) \leqslant f(\mathbf{x}) + \overline{\mathbf{u}}'\mathbf{g}(\overline{\mathbf{x}}) + \overline{\mathbf{v}}'\mathbf{h}(\overline{\mathbf{x}}) = \Phi(\mathbf{x}, \overline{\mathbf{u}}, \overline{\mathbf{v}})$$

for $\mathbf{x} \in X$.
Since $\overline{\mathbf{u}}'\mathbf{g}(\overline{\mathbf{x}}) = \overline{\mathbf{v}}'\mathbf{h}(\overline{\mathbf{x}}) = 0$,

$$\Phi(\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\mathbf{v}}) = f(\overline{\mathbf{x}}) + \overline{\mathbf{u}}'\mathbf{g}(\overline{\mathbf{x}}) + \overline{\mathbf{v}}'\mathbf{h}(\overline{\mathbf{x}}) \leqslant \Phi(\mathbf{x}, \overline{\mathbf{u}}, \overline{\mathbf{v}})$$

for $\mathbf{x} \in X$ (the second inequality of 3). The first inequality holds because $\overline{\mathbf{u}}'\mathbf{g}(\overline{\mathbf{x}}) = 0$, $\mathbf{h}(\overline{\mathbf{x}}) = \mathbf{0}$, $\mathbf{g}(\overline{\mathbf{x}}) \leqslant \mathbf{0}$ and $\mathbf{u} \geq \mathbf{0}$.

KT Conditions and Saddle Points

Theorem

Let $S = \{\mathbf{x} \in X \mid \mathbf{g}(\mathbf{x}) \leqslant \mathbf{0}, \mathbf{h}(\mathbf{x}) = \mathbf{0}\}$ *and consider the primal problem* $\mathcal{O}(f, \mathbf{g}, \mathbf{h})$. *Suppose that* $\overline{\mathbf{x}} \in S$ *satisfies the KTC, that is, there is* $\overline{\mathbf{u}} \geqslant \mathbf{0}$ *and* $\overline{\mathbf{v}}$ *such that*

$$\begin{aligned} (\nabla f)(\overline{\mathbf{x}}) + \overline{\mathbf{u}}'(\nabla \mathbf{g})(\overline{\mathbf{x}}) + \overline{\mathbf{v}}'(\nabla \mathbf{h})(\overline{\mathbf{x}}) &= \mathbf{0} \qquad (4) \\ \overline{\mathbf{u}}'\mathbf{g}(\overline{\mathbf{x}}) &= 0. \qquad (5) \end{aligned}$$

Suppose that f, g_i *for* $i \in I_{\overline{\mathbf{x}}}$ *are convex,* $v_j \neq 0$, *and* $\mathbf{h}(\mathbf{x}) = A\mathbf{x} - \mathbf{b}$. *Then* $\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\mathbf{v}}$ *satisfy the saddle point conditions*

$$\Phi(\overline{\mathbf{x}}, \mathbf{u}, \mathbf{v}) \leqslant \Phi(\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\mathbf{v}}) \leqslant \Phi(\mathbf{x}, \overline{\mathbf{u}}, \overline{\mathbf{v}})$$

for all $\mathbf{x} \in X$ *and all* $\mathbf{u}, \mathbf{v}$ *with* $\mathbf{u} \geqslant \mathbf{0}$, *where*

$$\Phi(\mathbf{x}, \mathbf{u}, \mathbf{v}) = f(\mathbf{x}) + \mathbf{u}'\mathbf{g}(\mathbf{x}) + \mathbf{v}'h(\mathbf{x}).$$

Proof

Suppose that $\bar{\mathbf{x}} \in S$ satisfies the KTC, that is, there is $\bar{\mathbf{u}} \geqslant \mathbf{0}$ and $\bar{\mathbf{v}}$ such that

$$\begin{aligned}(\nabla f)(\bar{\mathbf{x}}) + \bar{\mathbf{u}}'(\nabla \mathbf{g})(\bar{\mathbf{x}}) + \bar{\mathbf{v}}'(\nabla \mathbf{h})(\bar{\mathbf{x}}) &= \mathbf{0} && (6)\\ \bar{\mathbf{u}}'\mathbf{g}(\bar{\mathbf{x}}) &= 0.. && (7)\end{aligned}$$

Since f and g_i are convex, we have

$$\begin{aligned} f(\mathbf{x}) &\geqslant f(\bar{\mathbf{x}}) + (\nabla f)'(\mathbf{x} - \bar{\mathbf{x}}),\\ g_i(\mathbf{x}) &\geqslant g_i(\bar{\mathbf{x}}) + (\nabla g_i)'(\mathbf{x} - \bar{\mathbf{x}}) \text{ for } i \in I_{\bar{\mathbf{x}}}\\ h_j(\mathbf{x}) &= h_j(\bar{\mathbf{x}}) + (\nabla h_j)(\bar{\mathbf{x}})'(\mathbf{x} - \bar{\mathbf{x}}) \text{ for } 1 \leqslant j \leqslant \ell, (\bar{\mathbf{v}})_j \neq 0,\end{aligned}$$

for $\mathbf{x} \in X$.

Proof

Multiplying and adding the equalities:

$$\begin{array}{rcll} f(\mathbf{x}) & \geqslant & f(\overline{\mathbf{x}}) + (\nabla f)'(\mathbf{x} - \overline{\mathbf{x}}), & \\ g_i(\mathbf{x}) & \geqslant & g_i(\overline{\mathbf{x}}) + (\nabla g_i)'(\mathbf{x} - \overline{\mathbf{x}}) \text{ for } i \in I_{\overline{\mathbf{x}}} & \text{by } (\overline{\mathbf{u}})_i \\ h_j(\mathbf{x}) & = & h_j(\overline{\mathbf{x}}) + (\nabla h_j)(\overline{\mathbf{x}})'(\mathbf{x} - \overline{\mathbf{x}}) \text{ for } 1 \leqslant j \leqslant \ell, (\overline{\mathbf{v}})_j \neq 0, & \text{by } (\overline{\mathbf{v}})_j, \end{array}$$

and taking into account $\overline{\mathbf{u}}'\mathbf{g}(\overline{\mathbf{x}}) = 0$, it follows that $\Phi(\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\mathbf{v}}) \leqslant \Phi(\mathbf{x}, \overline{\mathbf{u}}, \overline{\mathbf{v}})$. Since $\mathbf{g}(\overline{\mathbf{x}}) \leqslant \mathbf{0}$, $\mathbf{g}(\overline{\mathbf{x}}) = \mathbf{0}$, and $\overline{\mathbf{u}}'\mathbf{g}(\overline{\mathbf{x}}) = 0$, it follows that

$$\Phi(\overline{\mathbf{x}}, \mathbf{u}, \mathbf{v}) \leqslant \Phi(\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\mathbf{v}}).$$

Thus, $\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\mathbf{v}}$ satisfy the saddle point conditions.

Reciprocal Theorem

Theorem

If $\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\mathbf{v}}$ *with* $\overline{\mathbf{x}} \in \mathbf{I}(X)$ *and* $\mathbf{u} \geqslant \mathbf{0}$ *satisfy the saddle point conditions, then* $\overline{\mathbf{x}}$ *is feasible to the primal problem, and* $\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\mathbf{v}}$ *satisfies the conditions 6 and 7.*

Proof

Suppose that $\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\mathbf{v}}$ with $\overline{\mathbf{x}} \in \mathbf{I}(X)$ and $\overline{\mathbf{u}} \geqslant \mathbf{0}$ satisfy

$$\Phi(\overline{\mathbf{x}}, \mathbf{u}, \mathbf{v}) \leqslant \Phi(\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\mathbf{v}}) \leqslant \Phi(\mathbf{x}, \overline{\mathbf{u}}, \overline{\mathbf{v}})$$

Since $\Phi(\overline{\mathbf{x}}, \mathbf{u}, \mathbf{v}) \leqslant \Phi(\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\mathbf{v}})$ for all $\mathbf{u} \geqslant \mathbf{0}$ and all $\mathbf{v}$, it follows that

$$\mathbf{g}(\overline{\mathbf{x}}) \leqslant \mathbf{0}, \mathbf{h}(\overline{\mathbf{x}}) = \mathbf{0}, \overline{\mathbf{u}}'\mathbf{g}(\overline{\mathbf{x}}) = 0.$$

Thus, $\overline{\mathbf{x}}$ is feasible to the primal. Since

$$\Phi(\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\mathbf{v}}) \leqslant \Phi(\mathbf{x}, \overline{\mathbf{u}}, \overline{\mathbf{v}}),$$

it follows that $\overline{\mathbf{x}}$ minimizes $\Phi(\mathbf{x}, \overline{\mathbf{u}}, \overline{\mathbf{v}})$ subject to $\mathbf{x} \in X$. Since $\overline{\mathbf{x}} \in \mathbf{I}(X)$, $\nabla_{\mathbf{x}}\Phi(\overline{\mathbf{x}}, \overline{\mathbf{u}}, \overline{\mathbf{v}}) = \mathbf{0}$, that is $(\nabla f)(\overline{\mathbf{x}}) + (\nabla \mathbf{g})(\overline{\mathbf{x}})\overline{\mathbf{u}} + (\nabla \mathbf{h})(\overline{\mathbf{x}})\overline{\mathbf{v}} = \mathbf{0}$.