Lagrangian Duality

José De Doná

September 2004

The University of Newcastle Australia

Centre of Complex Dynamic Systems and Control

Outline

Lagrangian Duality

- Given a nonlinear programming problem, known as the *primal problem*, there exists another nonlinear programming problem, closely related to it, that receives the name of the *Lagrangian dual problem*.

- Under certain convexity assumptions and suitable constraint qualifications, the primal and dual problems have equal optimal objective values.

The Primal Problem

Consider the following nonlinear programming problem:

Primal Problem P

$$\begin{aligned}
&\text{minimise } f(x), \qquad (1)\\
&\text{subject to:}\\
&g_i(x) \le 0 \quad \text{for } i = 1, \ldots, m,\\
&h_i(x) = 0 \quad \text{for } i = 1, \ldots, \ell,\\
&x \in X.
\end{aligned}$$

Then the *Lagrangian dual problem* is defined as the following nonlinear programming problem.

Lagrangian Dual Problem D

$$\begin{aligned} &\text{maximise } \theta(u, v), \qquad (2)\\ &\text{subject to:}\\ &u \geq 0, \end{aligned}$$

where,

$$\theta(u, v) = \inf\{f(x) + \sum_{i=1}^{m} u_i g_i(x) + \sum_{i=1}^{\ell} v_i h_i(x) : x \in X\}, \qquad (3)$$

is the *Lagrangian dual function*.

The Dual Problem

- In the dual problem (2)–(3), the vectors u and v have as their components the Lagrange multipliers u_i for $i = 1, \ldots, m$, and v_i for $i = 1, \ldots, \ell$.

- Note that the Lagrange multipliers u_i, corresponding to the inequality constraints $g_i(x) \leq 0$, are restricted to be nonnegative, whereas the Lagrange multipliers v_i, corresponding to the equality constraints $h_i(x) = 0$, are unrestricted in sign.

- Given the primal problem P (1), several Lagrangian dual problems D of the form of (2)–(3) can be devised, depending on which constraints are handled as $g_i(x) \leq 0$ and $h_i(x) = 0$, and which constraints are handled by the set X. (An appropriate selection of the set X must be made, depending on the nature of the problem.)

Consider the following primal problem P:

Primal Problem P

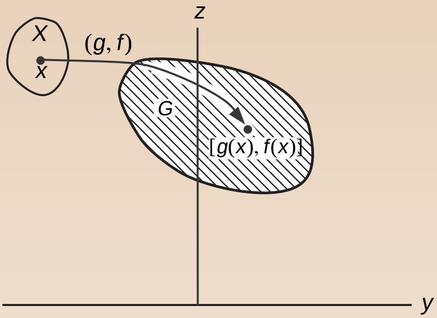
minimise $f(x)$,

subject to:

$g(x) \leq 0$,

$x \in X$,

where $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}$.



Define the following set in $\mathbb{R}^2$:

$$G = \{(y, z) : y = g(x), z = f(x) \text{ for some } x \in X\},$$

that is, G is the image of X under the (g, f) map.

Geometric Interpretation

$$G = \{(y, z) : y = g(x), z = f(x) \text{ for some } x \in X\},$$

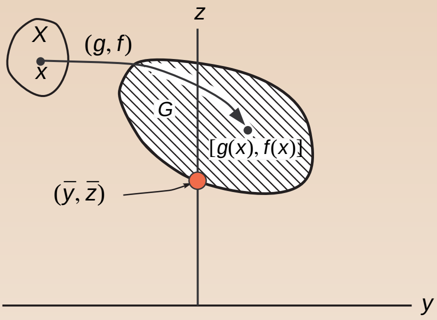
Primal Problem P

minimise $f(x)$,

subject to:

$g(x) \leq 0$,

$x \in X$.



Then, the primal problem consists in finding a point in G with $y \leq 0$ that has minimum ordinate z.

Obviously this point is $(\bar{y}, \bar{z})$.

Lagrangian Dual Problem D

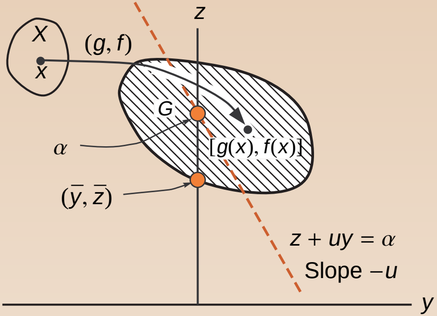
maximise $\theta(u)$,

subject to:

$u \geq 0$,

where (*Lagrangian dual subproblem*):

$\theta(u) = \inf\{f(x) + ug(x) : x \in X\}$.



Given $u \geq 0$, the Lagrangian dual subproblem is equivalent to minimise $z + uy$ over points (y, z) in G. Note that $z + uy = \alpha$ is the equation of a straight line with slope $-u$ that intercepts the z-axis at α.

Geometric Interpretation

Lagrangian Dual Problem D

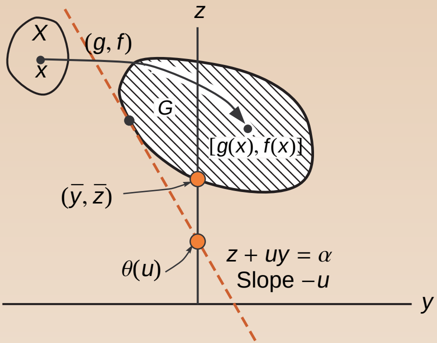
maximise $\theta(u)$,

subject to:

$u \geq 0$,

where (*Lagrangian dual subproblem*):

$\theta(u) = \inf\{f(x) + ug(x) : x \in X\}$.



In order to minimise $z + uy$ over G we need to move the line $z + uy = \alpha$ parallel to itself as far down as possible, whilst it remains in contact with G. The last intercept on the z-axis thus obtained is the value of $\theta(u)$ corresponding to the given $u \geq 0$.

Lagrangian Dual Problem D

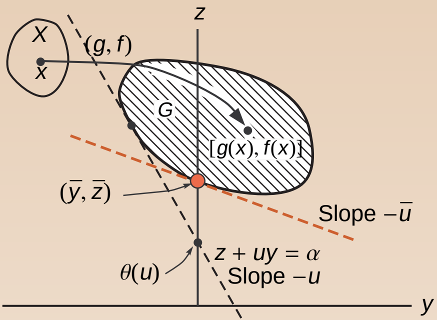
maximise $\theta(u)$,

subject to:

$u \geq 0$,

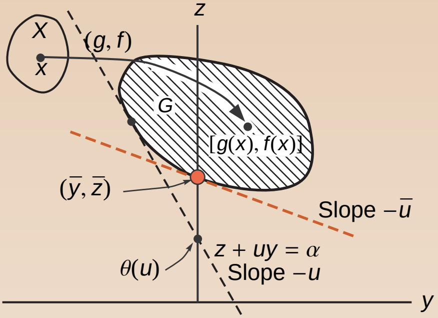
where (*Lagrangian dual subproblem*):

$\theta(u) = \inf\{f(x)+ug(x) : x \in X\}$.



Finally, to solve the dual problem, we have to find the line with slope $-u$ ($u \geq 0$) such that the last intercept on the z-axis, $\theta(u)$, is maximal. Such a line has slope $-\bar{u}$ and supports the set G at the point $(\bar{y}, \bar{z})$. Thus, the solution to the dual problem is $\bar{u}$, and the optimal dual objective value is $\bar{z}$.

Geometric Interpretation



- The solution of the Primal problem is $\bar{z}$, and the solution of the Dual problem is also $\bar{z}$.
- It can be seen that, in the example illustrated, the optimal primal and dual objective values are equal. In such cases, it is said that there is no *duality gap* (strong duality).

Weak Duality

The following result shows that the objective value of any feasible solution to the dual problem constitutes a lower bound for the objective value of any feasible solution to the primal problem.

Theorem (Weak Duality Theorem)

Consider the primal problem P *given by* (1) *and its Lagrangian dual problem* D *given by* (2). *Let* x *be a feasible solution to* P; *that is,* $x \in X$, $g(x) \leq 0$, *and* $h(x) = 0$. *Also, let* (u, v) *be a feasible solution to* D; *that is,* $u \geq 0$. *Then:*

$$f(x) \geq \theta(u, v).$$

$\circ$

Weak Duality

Proof.

We use the definition of θ given in (3), and the facts that $x \in X$, $u \geq 0$, $g(x) \leq 0$ and $h(x) = 0$. We then have

$$\begin{aligned}\theta(u, v) &= \inf\{f(\tilde{x}) + u^\intercal g(\tilde{x}) + v^\intercal h(\tilde{x}) : \tilde{x} \in X\} \\ &\leq f(x) + u^\intercal g(x) + v^\intercal h(x) \leq f(x),\end{aligned}$$

and the result follows. $\square$

Weak Duality

We then have, as a corollary of the previous theorem, the following result.

Corollary

$$\inf\{f(x) : x \in X, g(x) \leq 0, h(x) = 0\} \geq \sup\{\theta(u, v) : u \geq 0\}.$$

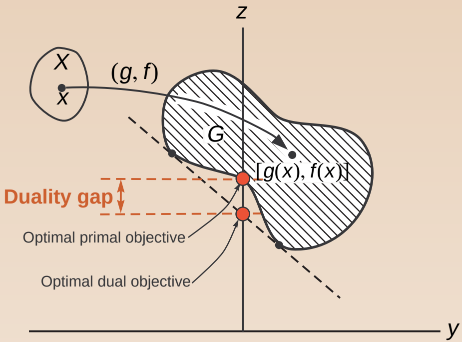
○

Note from the corollary that the optimal objective value of the primal problem is greater than or equal to the optimal objective value of the dual problem.

If the inequality holds as a *strict* inequality, then it is said that there exists a *duality gap*.

Weak Duality

The figure shows an example of the geometric interpretation of the primal and dual problems.



Notice that, in the case shown in the figure, there exists a duality gap due to the nonconvexity of the set G.

We will see, in the **Strong Duality Theorem**, that if some suitable convexity conditions are satisfied, then there is no duality gap between the primal and dual optimisation problems.

Strong Duality

Before stating the conditions that guarantee the absence of a duality gap, we need the following result.

Lemma

Let X be a nonempty convex set in $\mathbb{R}^n$. Let $\alpha : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ be (componentwise) convex, and $h : \mathbb{R}^n \to \mathbb{R}^\ell$ be affine (that is, assume h is of the form $h(x) = Ax - b$). Also, let u_0 be a scalar, $u \in \mathbb{R}^m$ and $v \in \mathbb{R}^\ell$. Consider the following two systems:

System 1: $\alpha(x) < 0, \quad g(x) \leq 0, \quad h(x) = 0$ *for some* $x \in X$.

System 2: $u_0\alpha(x) + u^\top g(x) + v^\top h(x) \geq 0$ *for some* $(u_0, u, v) \neq (0, 0, 0)$, $(u_0, u) \geq (0, 0)$ *and for all* $x \in X$.

If System 1 has no solution x, then System 2 has a solution (u_0, u, v). Conversely, if System 2 has a solution (u_0, u, v) with $u_0 > 0$, then System 1 has no solution.

Proof of the Lemma

Outline of the proof:

Assume first that

System 1: $\alpha(x) < 0, \quad g(x) \leq 0, \quad h(x) = 0 \quad$ for some $x \in X$,

has no solution.

Define the set:

$$S = \{(p, q, r) : p > \alpha(x), q \geq g(x), r = h(x) \quad \text{for some } x \in X\}.$$

The set S is convex, since X, α and g are convex and h is affine. Since System 1 has no solution, we have that $(0, 0, 0) \notin S$.

Proof of the Lemma (Ctd.)

Example

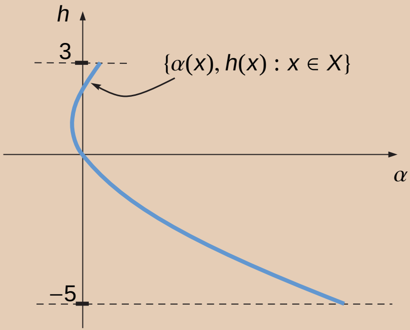
Consider the functions:

$$\alpha(x) = (x-1)^2 - \tfrac{1}{4},$$

$$h(x) = 2x - 1,$$

and the set

$$X = \{x \in \mathbb{R} : |x| \le 2\}.$$



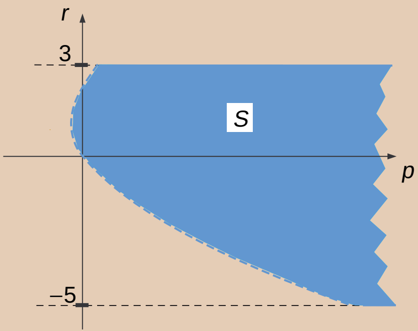
Proof of the Lemma (Ctd.)

Example (Ctd.)

$$\alpha(x) = (x-1)^2 - \tfrac{1}{4},$$

$$h(x) = 2x - 1,$$

$$X = \{x \in \mathbb{R} : |x| \le 2\}.$$



$$S = \{(p, r) : p > \alpha(x), r = h(x) \quad \textit{for some } x \in X\}$$

Proof of the Lemma (Ctd.)

Continuing with the proof of the Lemma, we have the convex set:

$$S = \{(p, q, r) : p > \alpha(x), q \geq g(x), r = h(x) \quad \text{for some } x \in X\},$$

and that $(0, 0, 0) \notin S$.

Recall the following corollary of the **Supporting Hyperplane Theorem**:

Corollary

Let S be a nonempty convex set in $\mathbb{R}^n$ and $\bar{x} \notin \operatorname{int} S$. Then there is a nonzero vector p such that $p^\mathsf{T}(x - \bar{x}) \leq 0$ for each $x \in \operatorname{cl} S$.

○

Proof of the Lemma (Ctd.)

We then have, from the above corollary, that there exists a nonzero vector (u_0, u, v) such that

$$(u_0, u, v)^\intercal[(p, q, r) - (0, 0, 0)] = u_0 p + u^\intercal q + v^\intercal r \geq 0, \qquad (4)$$

for each $(p, q, r) \in \mathrm{cl}\, S$.

Now, fix an $x \in X$. Noticing, from the definition of S, that p and q can be made arbitrarily large, we have that in order to satisfy (4), we must have $u_0 \geq 0$ and $u \geq 0$.

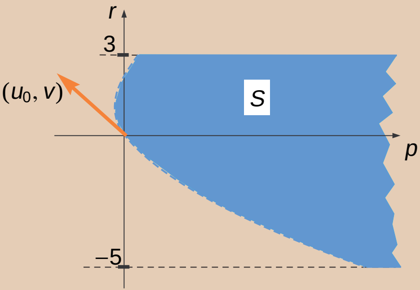
Proof of the Lemma (Ctd.)

Example (Ctd.)

$$\alpha(x) = (x-1)^2 - \tfrac{1}{4},$$

$$h(x) = 2x - 1,$$

$$X = \{x \in \mathbb{R} : |x| \leq 2\}.$$



We can see that u_0 cannot be $u_0 < 0$ and satisfy:

$$(u_0, v)^{\mathsf{T}}[(p, r) - (0, 0)] = (u_0, v)^{\mathsf{T}}(p, r) = u_0 p + v^{\mathsf{T}} r \geq 0,$$

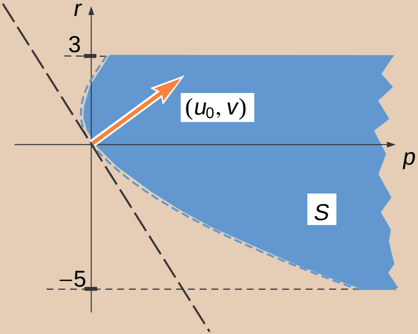
for each $(p, q, r) \in \mathrm{cl}\, S$.

Example (Ctd.)

$\alpha(x) = (x - 1)^2 - \frac{1}{4}$,

$h(x) = 2x - 1$,

$X = \{x \in \mathbb{R} : |x| \leq 2\}$.



We conclude that $u_0 \geq 0$ *and*

$$(u_0, v)^{\mathsf{T}}[(p, r) - (0, 0)] = (u_0, v)^{\mathsf{T}}(p, r) = u_0 p + v^{\mathsf{T}} r \geq 0,$$

for each $(p, q, r) \in \operatorname{cl} S$.

Proof of the Lemma (Ctd.)

We have that there exists a nonzero vector (u_0, u, v) with $(u_0, u) \geq (0, 0)$ such that

$$(u_0, u, v)^\mathsf{T}[(p, q, r) - (0, 0, 0)] = u_0 p + u^\mathsf{T} q + v^\mathsf{T} r \geq 0,$$

for each $(p, q, r) \in \mathrm{cl}\, S$.

Also, note that $[\alpha(x), g(x), h(x)] \in \mathrm{cl}\, S$ and we have from the above inequality that

$$u_0 \alpha(x) + u^\mathsf{T} g(x) + v^\mathsf{T} h(x) \geq 0.$$

Since the above inequality is true for each $x \in X$, we conclude that

System 2: $u_0 \alpha(x) + u^\mathsf{T} g(x) + v^\mathsf{T} h(x) \geq 0$ for some $(u_0, u, v) \neq (0, 0, 0)$, $(u_0, u) \geq (0, 0)$ and for all $x \in X$.

has a solution.

Proof of the Lemma (Ctd.)

To prove the converse, assume that

System 2: $u_0\alpha(x) + u^\intercal g(x) + v^\intercal h(x) \geq 0$ for some $(u_0, u, v) \neq (0, 0, 0)$, $(u_0, u) \geq (0, 0)$ and for all $x \in X$,

has a solution (u_0, u, v) such that $u_0 > 0$.

Suppose that $x \in X$ is such that $g(x) \leq 0$ and $h(x) = 0$.

From the previous inequality we conclude that $u_0\alpha(x) \geq -u^\intercal g(x) \geq 0$, since $u \geq 0$ and $g(x) \leq 0$. But, since $u_0 > 0$, we must then have that $\alpha(x) \geq 0$.

Hence,

System 1: $\alpha(x) < 0, \quad g(x) \leq 0, \quad h(x) = 0$ for some $x \in X$.

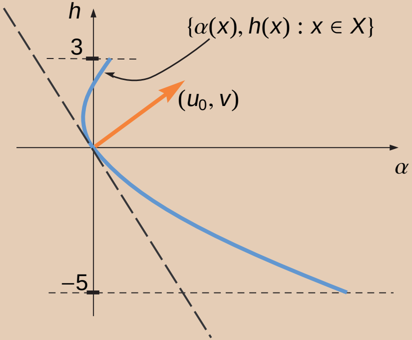
has no solution and this completes the proof. $\square$

Example (Ctd.)

$$\alpha(x) = (x-1)^2 - \tfrac{1}{4},$$

$$h(x) = 2x - 1,$$

$$X = \{x \in \mathbb{R} : |x| \le 2\}.$$



If $\Big\{$ ***System 2:*** $u_0\alpha(x) + v^{\mathsf{T}}h(x) \ge 0$ *for some* $(u_0, v) \neq (0,0)$, $u_0 \ge 0$ *and for all* $x \in X$ $\Big\}$, *has a solution such that* $u_0 > 0$, *and* $x \in X$ *is such that* $h(x) = 0$, *we can see that* $\alpha(x)$ *must be* $\alpha(x) \ge 0$, *and hence* ***System 1*** *has no solution.*

Strong Duality

The following result, known as the *strong duality theorem*, shows that, under suitable convexity assumptions and under a constraint qualification, there is no *duality gap* between the primal and dual optimal objective function values.

Theorem (Strong Duality Theorem)

Let X be a nonempty convex set in $\mathbb{R}^n$. Let $f : \mathbb{R}^n \to \mathbb{R}$ and $g : \mathbb{R}^n \to \mathbb{R}^m$ be convex, and $h : \mathbb{R}^n \to \mathbb{R}^\ell$ be affine. Suppose that the following constraint qualification is satisfied. There exists an $\hat{x} \in X$ such that $g(\hat{x}) < 0$ and $h(\hat{x}) = 0$, and $0 \in \operatorname{int} h(X)$, where $h(X) = \{h(x) : x \in X\}$. Then,

$$\inf\{f(x) : x \in X, g(x) \leq 0, h(x) = 0\} = \sup\{\theta(u, v) : u \geq 0\}, \tag{5}$$

where $\theta(u, v) = \inf\{f(x) + u^\intercal g(x) + v^\intercal h(x) : x \in X\}$. Furthermore, if the $\inf$ is finite, then $\sup\{\theta(u, v) : u \geq 0\}$ is achieved at $(\bar{u}, \bar{v})$ with $\bar{u} \geq 0$. If the $\inf$ is achieved at $\bar{x}$, then $\bar{u}^\intercal g(\bar{x}) = 0$.

Proof of the Strong Duality Theorem

Let $\gamma = \inf\{f(x) : x \in X, g(x) \leq 0, h(x) = 0\}$.

By assumption there exists a feasible solution $\hat{x}$ for the primal problem and hence $\gamma < \infty$.

If $\gamma = -\infty$, we then conclude from the corollary of the **Weak Duality Theorem** that $\sup\{\theta(u, v) : u \geq 0\} = -\infty$ and, hence, (5) is satisfied.

Thus, suppose that γ is finite, and consider the following system:

$$f(x) - \gamma < 0, \qquad g(x) \leq 0 \qquad h(x) = 0, \qquad \text{for some } x \in X.$$

By the definition of γ, this system has no solution. Hence, from the previous lemma, there exists a nonzero vector (u_0, u, v) with $(u_0, u) \geq (0, 0)$ such that

$$u_0[f(x) - \gamma] + u^\top g(x) + v^\top h(x) \geq 0 \qquad \text{for all } x \in X. \qquad (6)$$

Proof of the Strong Duality Theorem (Ctd.)

We will next show that $u_0 > 0$. Suppose, by contradiction that $u_0 = 0$.

By assumption, there exists an $\hat{x} \in X$ such that $g(\hat{x}) < 0$ and $h(\hat{x}) = 0$. Substituting in (6) we obtain $u^\top g(\hat{x}) \geq 0$. But, since $g(\hat{x}) < 0$ and $u \geq 0$, $u^\top g(\hat{x}) \geq 0$ is only possible if $u = 0$.

From (6), $u_0 = 0$ and $u = 0$ imply that $v^\top h(x) \geq 0$ for all $x \in X$. But, since $0 \in \operatorname{int} h(X)$, we can choose an $x \in X$ such that $h(x) = -\lambda v$, where $\lambda > 0$. Therefore, $0 \leq v^\top h(x) = -\lambda \|v\|^2$, which implies that $v = 0$.

Thus, it has been shown that $u_0 = 0$ implies that $(u_0, u, v) = (0, 0, 0)$, which is a contradiction. We conclude, then, that $u_0 > 0$.

Proof of the Strong Duality Theorem (Ctd.)

Dividing (6) by u_0 and denoting $\bar{u} = u/u_0$ and $\bar{v} = v/u_0$, we obtain

$$f(x) + \bar{u}^\intercal g(x) + \bar{v}^\intercal h(x) \geq \gamma \qquad \text{for all } x \in X. \tag{7}$$

This implies that $\theta(\bar{u}, \bar{v}) = \inf\{f(x) + \bar{u}^\intercal g(x) + \bar{v}^\intercal h(x) : x \in X\} \geq \gamma$.

We then conclude, from the **Weak Duality Theorem**, that $\theta(\bar{u}, \bar{v}) = \gamma$. And, from the corollary of the **Weak Duality Theorem**, we conclude that $(\bar{u}, \bar{v})$ solves the dual problem.

Finally, to complete the proof, assume that $\bar{x}$ is an optimal solution to the primal problem; that is, $\bar{x} \in X$, $g(\bar{x}) \leq 0$, $h(\bar{x}) = 0$ and $f(\bar{x}) = \gamma$.

From (7), letting $x = \bar{x}$, we get $\bar{u}^\intercal g(\bar{x}) \geq 0$. Since $\bar{u} \geq 0$ and $g(\bar{x}) \leq 0$, we get $\bar{u}^\intercal g(\bar{x}) = 0$.

This completes the proof. □

Example of Strong Duality

Example

Consider the following optimisation problem:

Primal Problem P

$$\begin{aligned} &\textit{minimise } (x-1)^2, \\ &\textit{subject to:} \\ &2x - 1 = 0, \\ &x \in X = \{x \in \mathbb{R} : |x| \le 2\}. \end{aligned}$$

It is clear that the optimal value of the objective function is equal to $\left(\frac{1}{2} - 1\right)^2 = \frac{1}{4}$, *since the feasible set is the singleton* $\left\{\frac{1}{2}\right\}$.

Example of Strong Duality

Example (Ctd.)

Lagrangian Dual Problem D

maximise $\theta(v)$,

where the Lagrangian dual function is,

$\theta(v) = \inf\{(x-1)^2 + v(2x-1) : |x| \leq 2\}$.

Differentiating w.r.t. x and equating to zero, we get that the optimiser of the dual Lagrangian subproblem is $x^* = -v + 1$ (if $-1 \leq v \leq 3$).

Hence $\theta(v) = (-v + 1 - 1)^2 + v(-2v + 2 - 1) = -v^2 + v$.

Differentiating w.r.t. v and equating to zero, we get that the optimiser of the dual problem is $v^* = \dfrac{1}{2}$ and the optimal value of the dual problem is $-{v^*}^2 + v^* = \dfrac{1}{4}$. Thus, there is no duality gap.

Example of Strong Duality

Example (Ctd.)