IB Optimisation: Lecture 4

Mike Tehranchi

University of Cambridge

1 May 2020

UNIVERSITY OF CAMBRIDGE

Consider the problem

$$\text{minimise } f(x) \text{ subject to } g(x) = b, \; x \in X.$$

Let

$$L(x, \lambda) = f(x) + \lambda^\top (b - g(x))$$

be the Lagrangian.

The Lagrangian sufficiency theorem says a feasible x^* is optimal IF there exists λ^* such that

$$L(x^*, \lambda^*) \leq L(x, \lambda^*)$$

for all $x \in X$. In particular, we have

$$\inf_{x \in X,\ g(x)=b} f(x) = \inf_{x \in X} L(x, \lambda^*)$$

When MUST λ^* exist?

Theorem (Lagrangian necessity)

For $b \in \mathbb{R}^m$, let

$$\varphi(b) = \inf_{x \in X,\ g(x)=b} f(x).$$

If φ is convex and finite, then there exists a λ depending on b such that

$$\inf_{x \in X,\ g(x)=b} f(x) = \inf_{x \in X} L(x, \lambda).$$

Furthermore, if φ is differentiable, then $\lambda = D\varphi(b)$.

Definition

The function

$$\varphi(b) = \inf_{x \in X,\ g(x)=b} f(x).$$

is called the *value function* of the family of problems

$$P_b: \text{ minimise } f(x) \text{ subject to } g(x) = b,\ x \in X.$$

Proof. Suppose the value function φ is convex. Fix $b \in \mathbb{R}^m$. By the supporting hyperplane theorem, there exists a vector $\lambda \in \mathbb{R}^m$ such that

$$
\begin{aligned}
\varphi(b) &= \inf_{c \in \mathbb{R}^m} [\varphi(c) + \lambda^\top (b - c)] \\
&= \inf_{c \in \mathbb{R}^m} \inf_{x \in X, g(x) = c} [f(x) + \lambda^\top (b - c)] \\
&= \inf_{c \in \mathbb{R}^m} \inf_{x \in X, g(x) = c} [f(x) + \lambda^\top (b - g(x))] \\
&= \inf_{x \in X} L(x, \lambda)
\end{aligned}
$$

□

Remark. By adding slack variables, we have

$$\inf_{x \in X,\ g(x) \leq b} f(x) = \inf_{x \in X, z \geq 0, g(x)+z=b} f(x)$$

so the theorem also holds when the functional constraint is an inequality.

Theorem (Sufficient condition for the convexity of the value function)

Suppose that

1. *the set X is convex,*
2. *the objective function f is convex, and*
3. *the functional constraint is $g(x) \leq b$*
4. *and g_j is convex for all $1 \leq j \leq m$.*

Then φ is convex.

Proof. The $m = 1$ case is on the example sheet. The $m > 1$ case is almost exactly the same. $\square$

Consider a factory owner who makes n different products out of m raw materials.

- He needs to choose amount x_i to make of the i-th product for each $i = 1, \dots, n$.
- Given a vector of amounts $x = (x_1, \dots, x_n)^\top$ of products to manufacture, the factory requires the amount $g_j(x)$ of the j-th raw material for $j = 1, \dots, m$.
- The amount of the j-th raw material available is b_j.
- Only non-negative amounts of products can be produced.
- Given the amounts x of products, the profit earned is $f(x)$.

The factory owner then tries to solve the problem to

$$\text{maximise } f(x) \text{ subject to } g(x) \leq b, \quad x \geq 0.$$

Let $\varphi(b)$ be the maximised profit as a function of the amount of raw materials.

The factory owner could solve this problem by considering a fictional world where there is a market for the raw material. Suppose the unit price of raw material j is λ_j. The total profit is now

$$L(x, \lambda) = f(x) + \lambda^\top (b - g(x)).$$

In this world, the factory owner's problem is to

$$\text{maximise } L(x, \lambda) \text{ subject to } x \geq 0.$$

- For each price vector λ, the optimal amount of products is the vector $x(\lambda)$.
- Now find a price vector λ^* such that the $x^* = x(\lambda^*)$ is feasible, i.e. $g(x^*) \leq b$

Such a λ^* is called a *shadow price* vector of the raw materials.

Economic interpretation of supporting hyperplane

Suppose that a market for raw material sprung into existence.

The shadow price λ_j^* of the j-th raw material should be the highest price such that the factory owner would buy more of that raw material.

That is, the factory owner would buying the basket of $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_m)^\top$ of raw materials if

$$\varphi(b+\varepsilon) - \varphi(b) \geq \lambda^\top \varepsilon$$

When ε is small, the left-hand side is approximately

$$\varphi(b+\varepsilon) - \varphi(b) = D\varphi(b)^\top \varepsilon$$

If the price λ_j of the j-th raw material is greater than $\frac{\partial \varphi}{\partial b_j}$, the factory owner would not buy any.
The shadow price vector is then $D\varphi(b)$ as before.

Consistency check: Suppose

- the inequality constraint $g(x) \leq b$ implies the sign constraint $\lambda \geq 0$.
- if the j-th constraint is not tight, i.e $g_j(x^*) < b_j$ so the factory owner does not use all of the available j-th raw material then the shadow price $\frac{\partial \varphi}{\partial b_j} = 0$ as there is no extra profit in acquiring a little more.
- On the other hand, the j-th slack variable z_j is positive
- by complementary slackness we have $\lambda_j = 0$.

Economic motivation of duality

Consider the point of view of someone who sells raw materials to the factory, and then consumes the finished products. If the amount of finished products is x and the price of raw materials is λ, then the raw material seller's profit is

$$\begin{aligned}\text{profit of raw material seller } &= \lambda^\top(g(x) - b) - f(x)\\ &= -\text{ profit of factory owner}\end{aligned}$$

But x is chosen by the factory owner to maximise *his* profits, so the raw material seller problem is essentially to

$$\text{minimise } \sup_{x \geq 0}[f(x) + \lambda^\top(b - g(x))]$$

Definition

Consider the *primal* problem

$$P: \ \text{minimise } f(x) \text{ subject to } g(x) = b, \quad x \in X.$$

The Lagrangian is

$$L(x, \lambda) = f(x) + \lambda^\top (b - g(x)).$$

The set of feasible Lagrange multipliers is

$$\Lambda = \{\lambda \in \mathbb{R}^m : \inf_{x \in X} L(x, \lambda) > -\infty\}$$

Definition (Continued)

The *dual objective function* $h : \Lambda \to \mathbb{R}$ is defined by

$$h(\lambda) = \inf_{x \in X} L(x, \lambda).$$

The *dual* problem is defined to be

$$D: \text{ maximise } h(\lambda) \text{ subject to } \lambda \in \Lambda.$$

The set Λ is the set of *feasible solutions to the dual problem.*

Remarks.

- The dual problem for a problem with inequality constraints is formulated by first introducing slack variables.
- The dual problem to a maximisation problem is formed analogously. In particular, the dual problem in this case is a minimisation problem.

Theorem (Weak duality)

Let x be feasible for P and let λ be feasible for D. Then

$$h(\lambda) \leq f(x)$$

and in particular

$$\sup_{\lambda \in \Lambda} h(\lambda) \leq \inf_{x \in X, g(x) = b} f(x).$$

Proof: Let x and λ be feasible for their respective problems.

$$\begin{aligned} h(\lambda) &= \inf\{L(\xi,\lambda) : \ \xi \in X\} \\ &\leq L(x,\lambda) \text{ for all } x \in X \\ &= f(x) \text{ for all feasible } x \end{aligned}$$

since $L(x,\lambda) = f(x) + \lambda^\top(b - g(x)) = f(x)$ when x is feasible. $\square$

Remark. In light of the above weak duality result, the Lagrangian sufficiency theorem can be rephrased:

> *If x^* and λ^* are feasible for P and D respectively, and if $h(\lambda^*) = f(x^*)$, then x^* is optimal for P.*

Remark. The difference

$$\inf_{x \in X, g(x)=b} f(x) - \sup_{\lambda \in \Lambda} h(\lambda)$$

is called the duality gap. Weak duality says that the duality gap is non-negative in general. By Lagrangian necessity, if the value function of P is convex, then there exists λ^* such that

$$\inf_{x \in X, g(x)=b} f(x) = \inf_{x \in X} L(x, \lambda^*) = h(\lambda^*)$$

so the duality gap is zero. This is called *strong duality.*

Consider the primal problem

$$P: \text{ maximise } c^\top x \text{ subject to } Ax \leq b, \; x \geq 0$$

where A is a $m \times n$ matrix, $b \in \mathbb{R}^m$, and $c \in \mathbb{R}^n$.
This is an example of a *linear program*.

The dual problem is found as follows:

- Introduce slack variables

$$P: \text{ maximise } c^\top x \text{ subject to } Ax + z = b, \; x \geq 0, z \geq 0$$

- The Lagrangian is

$$\begin{aligned} L(x, z, \lambda) &= c^\top x + \lambda^\top (b - Ax - z) \\ &= b^\top \lambda + (c - A^\top \lambda)^\top x - \lambda^\top z \end{aligned}$$

- The set Λ of feasible solutions to the dual problem is then

$$\begin{aligned} \Lambda &= \{\lambda \in \mathbb{R}^m : \sup_{x \geq 0, z \geq 0} L(x, z, \lambda) < \infty\} \\ &= \{\lambda \in \mathbb{R}^m : A^\top \lambda \geq c, \lambda \geq 0, \} \end{aligned}$$

- The dual objective function

$$\sup_{x \geq 0, z \geq 0} L(x, z, \lambda) = b^\top \lambda \text{ for } \lambda \in \Lambda$$

- The dual problem is then the

$$D: \text{ minimise } b^\top \lambda \text{ subject to } A^\top \lambda \geq c, \; \lambda \geq 0$$