6 The Lagrangian Dual

6.1 Example: further use of the Lagrangian sufficiency theorem

Consider the problem

$$\begin{aligned}&\text{minimize } \frac{1}{1+x_1}+\frac{1}{2+x_2}\\ &\text{s.t. } x_1+x_2=b, \quad x_1,x_2\ge 0.\end{aligned}$$

Solution. We define $X=\{x : x\ge 0\}$ and the Lagrangian

$$\begin{aligned}L(x,\lambda) &= \frac{1}{1+x_1}+\frac{1}{2+x_2}-\lambda(x_1+x_2-b)\\ &= \left(\frac{1}{1+x_1}-\lambda x_1\right)+\left(\frac{1}{2+x_2}-\lambda x_2\right)+\lambda b.\end{aligned}$$

Note that

$$\left(\frac{1}{1+x_1}-\lambda x_1\right) \text{ and } \left(\frac{1}{2+x_2}-\lambda x_2\right)$$

do not have a finite minimum in $x\ge 0$ unless $\lambda\le 0$. So we take $\lambda\le 0$. Observe that in the range $x\ge 0$ a function of the form $\left(\frac{1}{c+x}-\lambda x\right)$ will have its minimum either at $x=0$, if this function is increasing at 0, or at the stationary point of the function, occurring where $x>0$, if the function is decreasing at 0. So the minimum occurs at

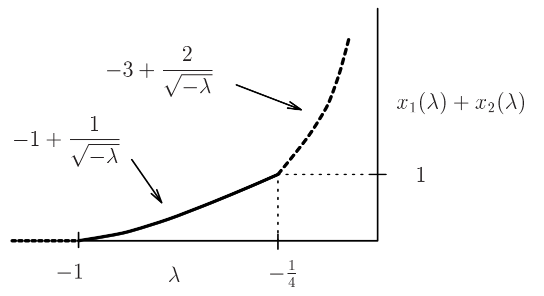
$$x=\begin{matrix}0\\ -c+\sqrt{-1/\lambda}\end{matrix} \quad \text{as} \quad \sqrt{-1/\lambda}\ \begin{matrix}\le\\ \ge\end{matrix}\ c.$$

Thus for $\sqrt{-1/\lambda}\le 1$ we have a minimum at $x_1=0$ and otherwise the minimum is at $x_1=-1+\sqrt{-1/\lambda}$. Similarly, if $\sqrt{-1/\lambda}\le 2$ we have a minimum at $x_2=0$ and otherwise the minimum is at $x_2=-2+\sqrt{-1/\lambda}$. So the set Y is $\{\lambda : \lambda\le 0\}$ and for $\lambda\in Y$, the minimum of $L(x,\lambda)$ in X is

$$\begin{array}{lllll} x_1=0, & x_2=0 & & & \le 1\\ x_1=-1+\sqrt{-1/\lambda}, & x_2=0 & & \text{as } \sqrt{-1/\lambda} & \in[1,2]\\ x_1=-1+\sqrt{-1/\lambda}, & x_2=-2+\sqrt{-1/\lambda} & & & \ge 2\end{array}$$

At first sight it appears that we don't know which values of x_1,x_2 to substitute into the constraint until we know λ, and we don't know λ until we substitute x_1,x_2 into the constraint. But notice that $x_1(\lambda)+x_2(\lambda)$ satisfies

$$x_1(\lambda)+x_2(\lambda)=\begin{cases}0 & \le -1\\ -1+1/\sqrt{-\lambda} \quad \text{as } \lambda & \in[-1,-1/4]\\ -3+2/\sqrt{-\lambda} & \in[-1/4,0]\end{cases}$$



So we can see that $x_1(\lambda)+x_2(\lambda)$ is an increasing and continuous function (although it is not differentiable at $\lambda=-1$ and at $\lambda=-1/4$). Thus (by the Intermediate Value Theorem) for any $b>0$ there will be a unique value of λ, say λ^*, for which $x_1(\lambda^*)+x_2(\lambda^*)=b$. This λ^* and corresponding x^* will satisfy the conditions of the theorem and so x^* is optimal. ■

Examples of this kind are fairly common.

6.2 The Lagrangian dual problem

We have already defined

$$Y=\{\lambda : \min_{x\in X} L(x,\lambda)>-\infty\}.$$

Now, for $\lambda\in Y$ define

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

The following lemma is almost as easy to prove as the sufficiency theorem.

Lemma 6.1 (weak duality) *For any feasible $x\in X_b$ and any $\lambda\in Y$*

$$L(\lambda)\le f(x).$$

Proof. For $x\in X_b$, $\lambda\in Y$,

$$f(x)=L(x,\lambda)\ge\min_{x\in X_b} L(x,\lambda)\ge\min_{x\in X} L(x,\lambda)=L(\lambda)$$

■

Thus, provided the set Y is non-empty, $L(\lambda)$, $\lambda\in Y$ provides a lower bound for the value of the objective function $f(x)$ at the optimum for P.

It seems reasonable to try and get this lower bound as good as possible, i.e., to consider the problem

$$\text{D: maximize } L(\lambda) \text{ subject to } \lambda\in Y,$$

equivalently,

$$\text{D: maximize}_{\lambda\in Y} \left\{\min_{x\in X} L(x,\lambda)\right\}.$$

This is known as the **dual problem**. (The original is the **primal problem**.)

In a moment we will return to linear problems and justify some of what we did in the first part of the course. For now, notice that the idea of a dual problem is quite general. For example, we can look again at the two examples we just studied.

In the example of Section 5.4 we had $Y = \{\lambda : \lambda_1 = -2, \lambda_2 < 0\}$ and that $\min_{x\in X} L(x,\lambda)$ occurred for $x(\lambda) = (3/2\lambda_2, 1/2\lambda_2, x_3)$. Thus

$$L(\lambda) = L(x(\lambda),\lambda) = \frac{10}{4\lambda_2} - 10 + 4\lambda_2.$$

The dual problem is thus

$$\underset{\lambda_2<0}{\text{maximize}} \left\{\frac{10}{4\lambda_2} - 10 + 4\lambda_2\right\}.$$

Note that the max occurs at $\lambda_2 = -\sqrt{5/8}$ and that the primal and dual have the same optimal value.

In the the example of Section 6.1, $Y = \{\lambda : \lambda \le 0\}$. By substituting the optimal value of x into $L(x,\lambda)$ we obtain

$$L(\lambda) = \begin{matrix} 3/2+\lambda b & & \le -1 \\ 1/2+2\sqrt{-\lambda}+(b+1)\lambda & \text{as } \lambda & \in[-1,-1/4] \\ 4\sqrt{-\lambda}+(b+3)\lambda & & \in[-1/4,0] \end{matrix}$$

We can solve the dual problem $\max L(\lambda)$ s.t. $\lambda \le 0$. The solution lies in $-1 \le \lambda \le -1/4$ if $0 \le b \le 1$ and in $-1/4 \le \lambda \le 0$ if $1 \le b$. You can confirm for all b that primal and dual here have the same optimal values.

6.3 The dual problem for LP

Construction of the dual problem for LP is relatively straightforward. Consider the primal problem P:

$$\begin{gathered} \text{maximize } c^\top x \\ \text{subject to } Ax \le b,\ x \ge 0 \\ \text{equivalently } Ax + z = b,\ x, z \ge 0. \end{gathered}$$

Write the Lagrangian

$$L(x,z,\lambda) = c^\top x - \lambda^\top(Ax+z-b) = (c^\top - \lambda^\top A)x - \lambda^\top z + \lambda^\top b.$$

As in the general case, we can find the set Y such that $\lambda \in Y$ implies $\max_{x,z\ge 0} L(x,z,\lambda)$ is finite, and for $\lambda \in Y$ we compute the minimum of $L(\lambda)$.

Consider the linear term $-\lambda^\top z$. If any coordinate $\lambda_i < 0$ we can make $-\lambda_i z_i$ as large as we like, by taking z_i large. So there is only a finite maximum in $z \ge 0$ if $\lambda_i \ge 0$ for all i.

Similarly, considering the term $(c^\top - \lambda^\top A)x$, this can be made as large as we like unless $(c^\top - \lambda^\top A)_i \le 0$ for all i. Thus

$$Y = \{\lambda : \lambda \ge 0,\ \lambda^\top A - c^\top \ge 0\}.$$

If we pick a $\lambda \in Y$ then $\max_{z\ge 0} -\lambda^\top z = 0$ (by choosing $z_i = 0$ if $\lambda_i > 0$ and any z_i if $\lambda_i = 0$) and also $\max_{x\ge 0}(c^\top - \lambda^\top A)x = 0$ similarly. Thus for $\lambda \in Y$, $L(\lambda) = \lambda^\top b$. So the dual problem for P is, as we have defined previously,

$$\text{D: } \min \lambda^\top b \quad \text{s.t. } A^\top \lambda \ge c,\ \lambda \ge 0.$$

Remark. Notice that in this section we have taken the primal to be a maximization problem and the dual to be a minimization problem. In Sections 5.2 and 6.2 these were the other way around. In general, the dual problem is always obtained by extremizing $L(\lambda)$ with respect to $\lambda \in Y$ in the opposite direction to that in which $L(x,\lambda)$ is extremized with respect to $x \in X$.

6.4 The weak duality theorem in the case of LP

We can now apply Lemma 6.1 directly to P and D to obtain one of the theorems which was first stated in Lecture 2.

Theorem 2.3 (weak duality) *If x is feasible for P (so $Ax \le b$, $x \ge 0$) and λ is feasible for D (so $\lambda \ge 0$, $A^\top\lambda \ge c$) then $c^\top x \le \lambda^\top b$.*

Since this is an important (though easy) result it may be worth writing out a proof for this particular case which does not appeal to the general Lemma 6.1. Naturally enough, the proof is very similar.

Proof. Write

$$L(x,z,\lambda) = c^\top x - \lambda^\top(Ax+z-b)$$

where $Ax + z = b$. Now for x and λ satisfying the conditions of the theorem,

$$c^\top x = L(x,z,\lambda) = (c^\top - \lambda^\top A)x - \lambda^\top z + \lambda^\top b \le \lambda^\top b.$$

■