7 Shadow Prices and Lagrangian Necessity

7.1 Sufficient conditions for optimality

Theorem 2.3 can be used to give a quick proof of the sufficient conditions for optimality.

Theorem 2.4 (sufficient conditions for optimality) *If x^*, z^* is feasible for P and λ^* is feasible for D and $(c^\top - \lambda^{*\top} A)x^* = \lambda^{*\top} z^* = 0$ (complementary slackness) then x^* is optimal for P and λ^* is optimal for D. Furthermore $c^\top x^* = \lambda^{*\top} b$.*

Proof. Write $L(x^*, z^*, \lambda^*) = c^\top x^* - \lambda^{*\top}(Ax^* + z^* - b)$. Now

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

But for all x feasible for P we have $c^\top x \leq \lambda^{*\top} b$ (by the weak duality theorem) and this implies that for all feasible x, $c^\top x \leq c^\top x^*$. So x^* is optimal for P. Similarly, λ^* is optimal for D (and the problems have the same optimums). ■

Remarks.

1. As an alternative to the above proof we could have just applied the Lagrangian sufficiency theorem directly to P using the λ^* given in the theorem. The proof comes out very similar.

2. The theorems above explain most of what we observed in Lecture 2. There we were looking at two problems, P and D, that were dual to one another. We paired basic solutions to P and D by insisting on complementary slackness. Then since our solution to P satisfied $Ax + z = b$ the same argument as in the proof above shows that we will obtain identical values of the objective function. (We don't use dual feasibility in the first line of the proof.) The theorem on sufficient conditions for optimality explains why it was that we only obtained a pair of solutions feasible for P and D at the optimum of both.

7.2 Shadow prices

We have mentioned before that dual variables or Lagrangian multipliers are sometimes known as shadow prices. First, an explanation of **shadow prices**.

Suppose you require an amount $b_1, \ldots, b_m$ of m different vitamins. There are n foodstuffs available. Let a_{ij} =amounts of vitamin i in one unit of foodstuff j, and suppose foodstuff j costs c_j per unit. Your problem is therefore to choose the amount

x_j of foodstuff j you buy to solve the LP

$$\min \sum_j c_j x_j$$
$$\text{subject to } \sum_j a_{ij} x_j \geq b_i, \text{ each } i$$
$$\text{and } x_j \geq 0 \text{ each } j.$$

Now suppose that a vitamin company decides to market m different vitamin pills (one for each vitamin) and sell them at price p_i per unit for vitamin i. Assuming you are prepared to switch entirely to a diet of vitamin pills, but that you are not prepared to pay more for an artificial carrot (vitamin equivalent) than a real one, the company has to maximize profit by choosing prices p_i to

$$\max \sum_i b_i p_i$$
$$\text{subject to } \sum_i a_{ij} p_i \leq c_j, \text{ each } j$$
$$\text{and } p_i \geq 0 \text{ each } i.$$

Note that this is the LP which is dual to your problem. The dual variable p_i is the price you are prepared to pay for a unit of vitamin i and is called a shadow price. By extension, dual variables are sometimes called shadow prices in problems where their interpretation as prices is very hard (or impossible) to see.

The dual variables tell us how the optimum value of our problem changes with changes in the right-hand side (b) of our functional constraints. This makes sense in the example given above. If you require an amount $b_i + \epsilon$ of vitamin i instead of an amount b_i you would expect the total cost of your foodstuff to change by an amount ϵp_i, where p_i is the value to you of a unit of vitamin i, even though in your problem you cannot buy vitamin i separately from the others.

Thus in the example in Section 5.4 we had constraints

$$x_1 + x_2 + x_3 = 5$$
$$x_1^2 + x_2^2 = 4$$

and obtained values of Lagrange multipliers of $\lambda_1^* = -2, \lambda_2^* = -\sqrt{5/8}$. If we replace the constraints by

$$x_1 + x_2 + x_3 = b_1$$
$$x_1^2 + x_2^2 = b_2$$

and write $\phi(b) =$ optimal value of the problem with $b = (b_1, b_2)^\top$ then you can check that

$$\left.\frac{\partial \phi}{\partial b_1}\right|_{b=(5,4)} = -2 \qquad \left.\frac{\partial \phi}{\partial b_2}\right|_{b=(5,4)} = -\sqrt{5/8}.$$

Let us explain a little more why this is true.

In linear problems we can use what we know about the optimal solutions to see how this works. Let us assume the primal problem

$$\begin{aligned} P(b): \qquad & \min c^\top x \\ \text{s.t. } & Ax + z = b, \ x \geq 0. \end{aligned}$$

has the optimal solution $\phi(b)$, depending on b. Consider two close together values of b, b' and b'' say, and suppose that optimal solutions have the same basic variables (so optimums occur with the same variables non-zero though the values of the variables will change slightly). The optimum still occurs at the same vertex of the feasible region though it moves slightly. Now consider the dual problem. This is

$$\begin{aligned} & \max \lambda^\top b \\ \text{s.t. } & A^\top \lambda \leq c, \ \lambda \leq 0. \end{aligned}$$

In the dual problem the feasible set does not depend on b, so the optimum of the dual will occur with the same basic variables and the same values of the dual variables λ. But the value of the optimum dual objective function is $\lambda^\top b'$ in one case and $\lambda^\top b''$ in the other and we have seen that the primal and dual have the same solutions. Hence

$$\phi(b') = \lambda^\top b' \text{ and } \phi(b'') = \lambda^\top b''$$

and the values of the dual variables λ give the rate of change of the objective value with b. The change is linear in this case.

Remarks.

1. You can also test this argument on the non-linear problems we have studied.

2. Recall that we made an argument that the objective row of the final tableau contains constants which give the rate of change of the optimum with b (based on the fact that the final tableau is the initial tableau + multiples of the original constraints). But we also observed that the objective row contains values of the dual variables at the optimum. That argument and the argument above are two versions of the same thing.

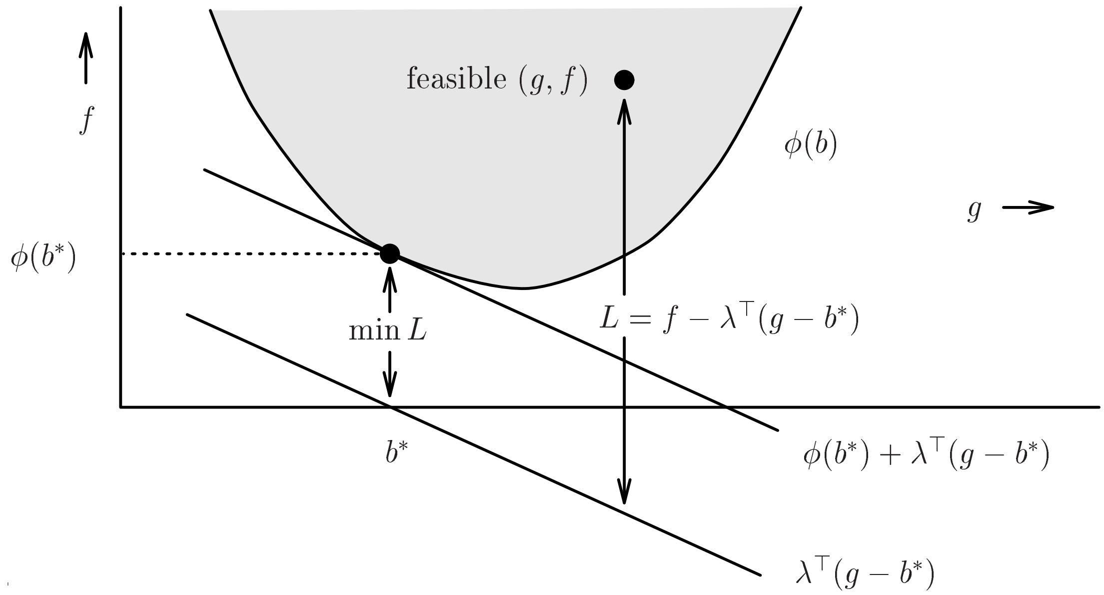
7.3 Lagrangian necessity

In all the examples we have studied we have been able to find λs such that we can apply the Lagrangian Sufficiency Theorem. Furthermore, we have observed that the primal and dual problems have the same optimum values. It is outside the scope of the course to establish conditions under which we expect these results to hold, but it seems helpful to give a brief summary.

Theorem 7.1 *If the functions f, g are convex, X is convex and x^* is an optimal solution to P, then there exist Lagrange multipliers $\lambda \in \mathbb{R}^m$ such that $L(x^*, \lambda) \leq L(x, \lambda)\ \forall x \in X$.*

In particular, Lagrange multipliers will always exist for LP problems provided the problem is bounded (and primal and dual problems will have the same optimal value).

Define $\phi(b)$ as $\min f(x)$, subject to $x \in X$ and $g(x) = b$. The proof of the theorem proceeds by first showing that $\phi(b)$ is convex. This implies that for each b^* there is a tangent hyperplane to $\phi(b)$ at b^* with the graph of $\phi(b)$ lying entirely above it. This uses the 'supporting hyperplane theorem' for convex sets, which is geometrically obvious, but some work to prove. Note the equation of the hyperplane will be $y = \phi(b^*) + \lambda^\top(b - b^*)$ for some multipliers λ. This λ can be shown to be the required vector of Lagrangian multipliers and the picture below gives some intuition of a geometric kind into why the Lagrange multipliers λ exist and why these λs give the rate of change of the optimum $\phi(b)$ with b.



7.4 Convexity of the feasible set in the case of LP

In Lecture 1 we advertised three general theorems for LPs. We should now prove these. We write the LP as $\min c^\top x$ s.t. $Ax = b$, $x \geq 0$.

Theorem 1.1 *The feasible set X_b is convex.*

Proof. Suppose $x \in X_b$ and $y \in X_b$. So $x, y \geq 0$ and $Ax = Ay = b$. Consider $z = \lambda x + (1 - \lambda)y$ for $0 \leq \lambda \leq 1$. Then $z_i = \lambda x_i + (1 - \lambda)y_i \geq 0$ for all i. So $z \geq 0$. Secondly,

$$Az = A(\lambda x + (1 - \lambda)y) = \lambda Ax + (1 - \lambda)Ay = \lambda b + (1 - \lambda)b = b.$$

So $z \in X_b$ and hence X_b is convex. ■