Lagrangian duality (11/22)

Let's now consider more general constrained optimization problems:

$$\begin{aligned} &\text{minimize} \quad f(x) \\ &\text{subject to} \quad g_i(x) \le 0, \quad 1 \le i \le m. \end{aligned} \qquad (*)$$

How to solve? One idea: let $I(x)$ be defined by:

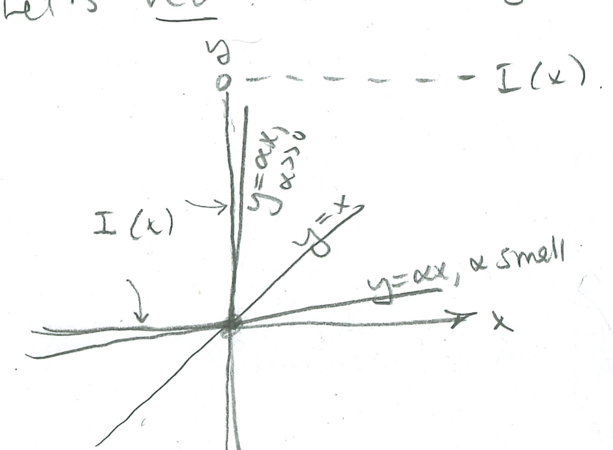
$$I(x) = \begin{cases} 0 & \text{if } x \le 0 \\ \infty & \text{else} \end{cases}$$

Then, if we solve:

$$\text{minimize} \quad f(x) + \sum_{i=1}^{m} I(g_i(x))$$

we have solved $(*)$, and vice versa. Of course, there are some obvious issues with $I(x)$... we remove the constraints to get an unconstrained optimization problem, but we haven't made the problem any easier to solve.

Let's relax $I(x)$, by replacing it with the function $x \mapsto \alpha x$:



For Large α, rejection from boundary, but for small α, good approximation of cost function in interior... so, a trade-off.

If $\alpha > 0$, then the penalty is in the right direction: the cost function is increased if we are outside the feasible set and reduced if we are inside the feasible set.

Let:

$$J(x) = f(x) + \sum_{i=1}^{m} I(g_i(x))$$

and let:

$$L(x, \lambda) = f(x) + \sum_{i=1}^{m} \lambda_i g_i(x).$$

<u>Claim</u>: $\max_{\lambda} L(x, \lambda) = J(x)$.

<u>Pf</u>: if x is feasible, then $g_i(x) \leq 0$ for $i = 1, \ldots, m$. Hence, to maximize L, set $\lambda = 0$. On the other hand, if x is infeasible, there exists some i such that $g_i(x) > 0$. Then, the maximization is unbounded, since we can just set $\lambda_i = \infty$.

Hence, for each constraint, it is clear that maximizing over λ effectively replaces "$\lambda_i g_i(x)$" with "$I(g_i(x))$", giving the result. ∎

How does this help?

Recall that solving (*) equivalent to solving: "$\min_x J(x)$".

But:

$$\min_x J(x) = \min_x \max_\lambda L(x,\lambda).$$

Let $g(\lambda) = \min_x L(x,\lambda)$. If:

$$\min_x \max_\lambda L(x,\lambda) = \max_\lambda \min_x L(x,\lambda) = \max_\lambda g(\lambda),$$

then we can solve $\max_\lambda g(\lambda)$ instead.

The function $g(\lambda)$ is called the <u>dual function</u> and $\max_\lambda g(\lambda)$ is called the <u>dual problem</u> (as opposed to the <u>primal problem</u>).

<u>Claim</u>: in general, we have: $\max_y \min_x f(x,y) \le \min_x \max_y f(x,y)$.

<u>Pf</u>: Let $x_* = \arg\min_x f(x,y)$, $y_* = \arg\max_y f(x,y)$. Then:

$$\min_x f(x,y) = f(x_*,y) \le f(x,y) \le f(x,y_*) = \max_y f(x,y).$$

Hence:

$$\min_x f(x,y) \le \max_y f(x,y) \quad \forall x,y.$$

Note that the LHS is constant independent of x, but the RHS is a function of x. So, the LHS provides a lower bound over all choices of x for the RHS, so it follows that:

$$\min_x f(x,y) \le \min_x \max_y f(x,y) \quad \forall y.$$

Likewise, this RHS is constant with respect to both x and y, and therefore gives an upper bound for this LHS, so it follows that:

$$\max_{y} \min_{x} f(x,y) \leq \min_{x} \max_{y} f(x,y) \quad \boxtimes$$

So, in general, we have:

$$\max_{\lambda} g(\lambda) \leq \min_{x} J(x),$$

but when does equality hold?

<u>Claim</u>: $\max_{y} \min_{x} f(x,y) = \min_{x} \max_{y} f(x)$ holds

if and only if there exists a point (x_*, y_*) for which the <u>saddle-point</u> condition holds:

$$f(x_*, y) \leq f(x_*, y_*) \leq f(x, y_*).$$

(That is: x_* is a minimizer when y is held fixed at y_*, and y_* is a maximizer when x is held fixed at x_*.)

<u>Pf</u>: ($\Rightarrow$) Assume $=$ holds. Then:

$$f(x_*, y) \leq \max_{y} f(x_*, y) = \max_{y} \min_{\tilde{x}} f(\tilde{x}, y) = f(x_*, y_*)$$
$$= \min_{x} \max_{y} f(x,y) = \min_{x} f(x, y_*) \leq f(x, y_*).$$

Hence, the saddle-point condition holds.

($\Leftarrow$) Suppose the saddle-point condition holds. Then:

$$LHS = \max_y f(x_*, y) \leq f(x_*, y_*) \leq \min_x f(x, y_*) = RHS$$

Similarly to before, it is clear that:

$$\min_x \max_y f(x,y) \leq \max_y f(x_*, y) = LHS,$$

and:

$$RHS = \min_x f(x, y_*) \leq \max_y \min_x f(x,y).$$

Hence:

$$\min_x \max_y f(x,y) \leq \max_y \min_x f(x,y).$$

But we already had:

$$\max_y \min_x f(x,y) \leq \min_x \max_y f(x,y).$$

Hence:

$$\max_y \min_x f(x,y) = \min_x \max_y f(x,y) \quad \blacksquare$$

When this happens, we say <u>strong duality</u> holds.

When strong duality holds, we can solve the dual problem instead of the primal problem and expect to get the same result.