Duality

Lagrangian

Consider general minimization problem

$$\begin{aligned} \min_x \quad & f(x) \\ \text{subject to} \quad & h_i(x) \le 0,\ i = 1, \dots m \\ & \ell_j(x) = 0,\ j = 1, \dots r \end{aligned}$$

Need not be convex, but of course we will pay special attention to convex case

We define the Lagrangian as

$$L(x, u, v) = f(x) + \sum_{i=1}^{m} u_i h_i(x) + \sum_{j=1}^{r} v_j \ell_j(x)$$

New variables $u \in \mathbb{R}^m, v \in \mathbb{R}^r$, with $u \ge 0$ (implicitly, we define $L(x, u, v) = -\infty$ for $u < 0$)

Important property: for any $u \geq 0$ and v,

$$f(x) \geq L(x, u, v) \quad \text{at each feasible } x$$

Why? For feasible x,

$$L(x, u, v) = f(x) + \sum_{i=1}^{m} u_i \underbrace{h_i(x)}_{\leq 0} + \sum_{j=1}^{r} v_j \underbrace{\ell_j(x)}_{=0} \leq f(x)$$

Lagrange dual function

Let C denote primal feasible set, $f^\star$ denote primal optimal value. Minimizing $L(x,u,v)$ over all x gives a lower bound:

$$f^\star \;\geq\; \min_{x \in C}\; L(x,u,v) \;\geq\; \min_x\; L(x,u,v) \;:=\; g(u,v)$$

We call $g(u,v)$ the Lagrange dual function, and it gives a lower bound on $f^\star$ for any $u \geq 0$ and v, called dual feasible u, v

Lagrange dual problem

Given primal problem

$$\begin{aligned} \min_x \quad & f(x) \\ \text{subject to} \quad & h_i(x) \leq 0, \ i = 1, \dots m \\ & \ell_j(x) = 0, \ j = 1, \dots r \end{aligned}$$

Our constructed dual function $g(u, v)$ satisfies $f^\star \geq g(u, v)$ for all $u \geq 0$ and v. Hence best lower bound is given by maximizing $g(u, v)$ over all dual feasible u, v, yielding Lagrange dual problem:

$$\begin{aligned} \max_{u,v} \quad & g(u, v) \\ \text{subject to} \quad & u \geq 0 \end{aligned}$$

Key property, called weak duality: if dual optimal value is $g^\star$, then

$$f^\star \geq g^\star$$

Note that this always holds (even if primal problem is nonconvex)

Another key property: the dual problem is a convex optimization problem (as written, it is a concave maximization problem)

Again, this is always true (even when primal problem is not convex)

By definition:

$$\begin{aligned} g(u,v) &= \min_x \; \Big\{ f(x) + \sum_{i=1}^m u_i h_i(x) + \sum_{j=1}^r v_j \ell_j(x) \Big\} \\ &= -\max_x \; \underbrace{\Big\{ -f(x) - \sum_{i=1}^m u_i h_i(x) - \sum_{j=1}^r v_j \ell_j(x) \Big\}}_{\text{pointwise maximum of convex functions in } (u,v)} \end{aligned}$$

I.e., g is concave in (u,v), and $u \geq 0$ is a convex constraint, hence dual problem is a concave maximization problem

Strong duality

Recall that we always have $f^\star \geq g^\star$ (weak duality). On the other hand, in some problems we have observed that actually

$$f^\star = g^\star$$

which is called strong duality

Slater's condition: if the primal is a convex problem (i.e., f and $h_1, \ldots h_m$ are convex, $\ell_1, \ldots \ell_r$ are affine), and there exists at least one strictly feasible $x \in \mathbb{R}^n$, meaning

$$h_1(x) < 0, \ldots h_m(x) < 0 \quad \text{and} \quad \ell_1(x) = 0, \ldots \ell_r(x) = 0$$

then strong duality holds

This is a pretty weak condition. An important refinement: strict inequalities only need to hold over functions h_i that are not affine

Duality gap

Given primal feasible x and dual feasible u, v, the quantity

$$f(x) - g(u, v)$$

is called the duality gap between x and u, v. Note that

$$f(x) - f^\star \le f(x) - g(u, v)$$

so if the duality gap is zero, then x is primal optimal (and similarly, u, v are dual optimal)

From an algorithmic viewpoint, provides a stopping criterion: if $f(x) - g(u, v) \le \epsilon$, then we are guaranteed that $f(x) - f^\star \le \epsilon$

Very useful, especially in conjunction with iterative methods ... more dual uses in coming lectures

Examples during class: linear and squared