

The Wolfe Dual

CS 246/446 Notes

1 Constrained Optimization

Consider a constrained optimization problem:

$$\begin{aligned} \min_x f(x) \\ \text{s.t. } g(x) = 0 \end{aligned}$$

At the solution, the gradient of the objective function f must be perpendicular to the constraint surface (feasible set) defined by $g(x) = 0$, so there exists a scalar Lagrange multiplier λ such that

$$\frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0$$

at the solution.

2 Inequalities

Consider an optimization problem with constraints specified as inequalities:

$$\begin{aligned} \min_x f(x) \\ \text{s.t. } g(x) \geq 0 \end{aligned}$$

If, at the solution, $g(x) = 0$, then as before there exists a λ such that

$$\frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0 \tag{1}$$

and furthermore $\lambda > 0$, otherwise we would be able to decrease $f(x)$ by moving in the direction $-\frac{\partial f}{\partial x}$ without leaving the feasible set defined by $g(x) \geq 0$.

If, on the other hand, at the solution $g(x) > 0$, then we must be at a maximum of $f(x)$, so $\frac{\partial f}{\partial x} = 0$ and

$$\frac{\partial f}{\partial x} - \lambda \frac{\partial g}{\partial x} = 0$$

with $\lambda = 0$. In either case, the following system of equations (known as the KKT conditions) holds:

$$\begin{aligned} \lambda g(x) &= 0 \\ \lambda &\geq 0 \\ g(x) &\geq 0 \end{aligned}$$

3 Convex Optimization

Suppose now that $f(x)$ is convex, and $g(x)$ is concave, and both are continuously differentiable. Define

$$L(x, \lambda) = f(x) - \lambda g(x)$$

and Equation 1 is equivalent to

$$\frac{\partial L}{\partial x} = 0$$

For any fixed $\lambda \geq 0$, L is convex in x , and has a unique minimum. For any fixed x , L is linear in λ .

Define

$$h(\lambda) = \min_x L(x, \lambda)$$

The minimum of a set of linear functions is concave, and has a maximum corresponding to the linear function with derivative of 0. Thus $h(\lambda)$ also has a unique maximum over $\lambda \geq 0$. Either the maximum of h occurs at $\lambda = 0$, in which case

$$h(0) = \min_x L(x, 0) = \min_x f(x)$$

and we are at the global minimum of f , or the maximum of h occurs at

$$\frac{\partial L}{\partial \lambda} = g(x) = 0$$

and we are on the boundary of the feasible set. In either case, the problem

$$\begin{aligned} \max_{\lambda} h(\lambda) \\ \text{s.t. } \lambda \geq 0 \end{aligned}$$

is equivalent to

$$\begin{aligned} \max_{\lambda, x} L(x, \lambda) \\ \text{s.t. } \lambda \geq 0 \\ \frac{\partial L}{\partial x} = 0 \end{aligned}$$

This is known as the dual problem, and its solution is also the solution to the original (primal) problem

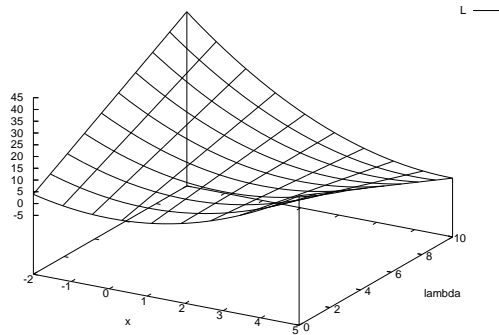
$$\begin{aligned} \min_x f(x) \\ \text{s.t. } g(x) = 0 \end{aligned}$$

4 An Example

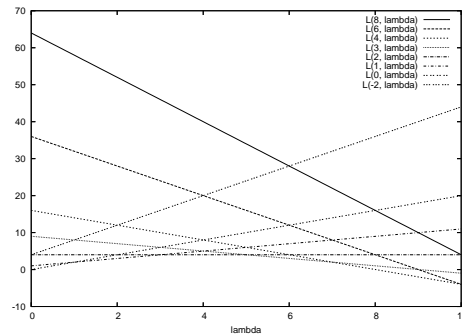
Minimize x^2 subject to $x \geq 2$.

$$\begin{aligned} L(x, \lambda) &= f(x) - \lambda g(x) \\ &= x^2 - \lambda(x - 2) \end{aligned}$$

The Lagrangian function L has a saddle shape:



Projecting onto the λ dimension, we see the concave function h formed from the minimum of linear functions $L(c, \lambda)$



To find h , set

$$\begin{aligned} \frac{\partial L}{\partial x} &= 0 \\ 2x - \lambda &= 0 \end{aligned}$$

and solve for x : $x = \lambda/2$. Substituting $x = \lambda/2$ into L gives $h(\lambda) = -\frac{1}{4}\lambda^2 - 2\lambda$. Setting $\frac{\partial h}{\partial \lambda} = 0$ yields $\lambda = 4$, which we see is the maximum of the concave shape in the figure. Substituting back into the original problem yields $x = 2$, a solution on the boundary of the constraint surface.