AM205: Constrained optimization using Lagrange multipliers

As discussed in the lectures, many practical optimization problems involve finding the minimum (or maximum) of some function over a set of parameters, subject to some constraints on those parameters. Mathematically, this can be expressed as finding the minimum of a function $f : \mathbb{R}^n \to \mathbb{R}$ subject to a number of constraints expressed as $g(x) = 0$, where $g : \mathbb{R}^n \to \mathbb{R}^m$ and $m \leq n$.

As described in the lectures, one convenient approach to solve optimization problems such as this is to introduce a vector of auxiliary variables $\lambda \in \mathbb{R}^m$. One can then introduce an augmented function,

$$\mathcal{L}(x, \lambda) = f(x) + \lambda^\mathsf{T} g(x). \tag{1}$$

One can show that the stationary points of $\mathcal{L}$ are the stationary points of f, subject to the constraints that $g(x) = 0$. The auxiliary variables λ are called the *Lagrange multipliers* and $\mathcal{L}$ is called the *Lagrangian function*. Lagrange multipliers can be used in computational optimization, but they are also useful for solving analytical optimization problems subject to constraints. Here, we consider a simple analytical example to examine how they work.

Example: constrained optimization of a cylinder's volume

Consider a cylinder of radius r and height h. Suppose we wish to find the minimum surface area of the cylinder, subject to the constraint that its volume is V. For this problem, we aim to minimize

$$f(r, h) = 2\pi r^2 + 2\pi r h = 2\pi r(r + h) \tag{2}$$

where the first term represents the curved cylindrical surface area and the second term represents the circular end caps. The constraint is

$$g(r, h) = \pi r^2 h - V. \tag{3}$$

Using the Lagrange multipliers approach, the Lagrangian function is

$$\mathcal{L}(r, h, \lambda) = 2\pi r(r + h) + \lambda(\pi r^2 h - V). \tag{4}$$

The stationary points of $\mathcal{L}$ are given by taking the first derivatives and setting them to zero. This gives

$$0 = \frac{\partial \mathcal{L}}{\partial r} = 4\pi r + 2\pi h + 2\lambda \pi r h, \tag{5}$$

$$0 = \frac{\partial \mathcal{L}}{\partial h} = 2\pi r + \lambda \pi r^2, \tag{6}$$

$$0 = \frac{\partial \mathcal{L}}{\partial \lambda} = \pi r^2 h - V. \tag{7}$$

Rearranging Eq. 6 shows that

$$\lambda = -\frac{2}{r} \tag{8}$$

and substituting this into Eq. 5 shows that

$$0 = 2r + h + (\lambda r)h = 2r + h - 2h = 2r - h \tag{9}$$

and hence $r = \frac{h}{2}$. Substituting into Eq. 7 gives

$$V = 2\pi r^3 \tag{10}$$

and hence

$$r = \sqrt[3]{\frac{V}{2\pi}}, \qquad h = \sqrt[3]{\frac{4V}{\pi}}. \tag{11}$$