1.23 Lagrange Multipliers

The maxima and minima of a function $f(x)$ of several variables $x_1, x_2, \ldots, x_n$ are among the points at which its gradient vanishes

$$\nabla f(x) = 0. \tag{1.221}$$

These are the stationary points of f.

Example 1.30 (Minimum)**.** For instance, if $f(x) = x_1^2 + 2x_2^2 + 3x_3^2$, then its minimum is at

$$\nabla f(x) = (2x_1, 4x_2, 6x_3) = 0 \tag{1.222}$$

that is, at $x_1 = x_2 = x_3 = 0$. □

But how do we find the extrema of $f(x)$ if x must satisfy a constraint? We use a Lagrange multiplier (Joseph-Louis Lagrange 1736–1813).

In the case of one constraint $c(x) = 0$, we no longer expect the gradient $\nabla f(x)$ to vanish, but its projection must vanish in those directions dx that preserve the constraint. So $dx \cdot \nabla f(x) = 0$ for all dx that make the dot product $dx \cdot \nabla c(x)$ vanish. This means that $\nabla f(x)$ and $\nabla c(x)$ must be parallel. So the extrema of $f(x)$ subject to the constraint $c(x) = 0$ satisfy two equations

$$\nabla f(x) = \lambda \, \nabla c(x) \quad \text{and} \quad c(x) = 0. \tag{1.223}$$

These equations define the extrema of the unconstrained function

$$L(x, \lambda) = f(x) - \lambda \, c(x) \tag{1.224}$$

of the $n + 1$ variables $x, \ldots, x_n, \lambda$

$$\nabla L(x, \lambda) = \nabla f(x) - \lambda \, \nabla c(x) = 0 \quad \text{and} \quad \frac{\partial L(x, \lambda)}{\partial \lambda} = \, - \, c(x) = 0. \tag{1.225}$$

The variable λ is a **Lagrange multiplier**.

In the case of k constraints $c_1(x) = 0, \ldots, c_k(x) = 0$, the projection of ∇f must vanish in those directions dx that preserve all the constraints. So $dx \cdot \nabla f(x) = 0$ for all dx that make all $dx \cdot \nabla c_j(x) = 0$ for $j = 1, \ldots, k$. The gradient ∇f will satisfy this requirement if it's a linear combination

$$\nabla f = \lambda_1 \, \nabla c_1 + \cdots + \lambda_k \, \nabla c_k \tag{1.226}$$

of the k gradients because then $dx \cdot \nabla f$ will vanish if $dx \cdot \nabla c_j = 0$ for $j = 1, \ldots, k$. The extrema also must satisfy the constraints

$$c_1(x) = 0, \ldots, c_k(x) = 0. \tag{1.227}$$

Equations (1.226 & 1.227) define the extrema of the unconstrained function

$$L(x, \lambda) = f(x) - \lambda_1 \, c_1(x) + \ldots \lambda_k \, c_k(x) \tag{1.228}$$

of the $n + k$ variables x and λ

$$\nabla L(x, \lambda) = \nabla f(x) - \lambda \, \nabla c_1(x) - \cdots - \lambda \, \nabla c_k(x) = 0 \tag{1.229}$$

and

$$\frac{\partial L(x, \lambda)}{\partial \lambda_j} = \, - \, c_j(x) = 0 \quad \text{for } j = 1, \ldots, k. \tag{1.230}$$

Example 1.31 (Constrained Extrema and Eigenvectors)**.** Suppose we want to find the extrema of a real, symmetric quadratic form $f(x) = x^{\mathsf{T}} A \, x$ subject to the constraint $c(x) = x \cdot x - 1$ which says that the vector x is of unit length. We form the function

$$L(x, \lambda) = x^{\mathsf{T}} A \, x - \lambda \, (x \cdot x - 1) \tag{1.231}$$

and since the matrix A is real and symmetric, we find its unconstrained extrema as

$$\nabla L(x, \lambda) = 2A \, x - 2\lambda \, x = 0 \quad \text{and} \quad x \cdot x = 1. \tag{1.232}$$

The extrema of $f(x) = x^{\mathsf{T}} A \, x$ subject to the constraint $c(x) = x \cdot x - 1$ are the normalized **eigenvectors**

$$A \, x = \lambda \, x \quad \text{and} \quad x \cdot x = 1. \tag{1.233}$$

of the real, symmetric matrix A. □

1.24 Eigenvectors

If a linear operator A maps a nonzero vector $|u\rangle$ into a multiple of itself

$$A|u\rangle = \lambda|u\rangle \tag{1.234}$$

then the vector $|u\rangle$ is an **eigenvector** of A with **eigenvalue** λ. (The German adjective *eigen* means special or proper.)

If the vectors $\{|k\rangle\}$ for $k = 1, \ldots, N$ form a basis for the vector space in which A acts, then we can write the identity operator for the space as