Math 16B – S06 – Supplementary Notes 3
The Lagrange Multiplier Method

The method of Lagrange multipliers is often effective in finding solutions of constrained extremum problems. In the two-variable version of such a problem, one is given a function $f(x, y)$, and one wishes to maximize it or minimize it under the constraint that another function $g(x, y)$ vanishes (i.e., one wishes to find a maximum or minimum of f on the level curve $g(x, y) = 0$). As explained in our textbook (where you will also find examples), Lagrange's method proceeds as follows. One introduces a third variable λ (traditionally called a Lagrange multiplier), and one defines a function $F(x, y, \lambda)$ of three variables by

$$F(x, y, \lambda) = f(x, y) + \lambda g(x, y).$$

The basic theorem underlying the method states that if $f(x, y)$ attains a maximum or a minimum at the point (a, b) under the constraint $g(x, y) = 0$, then there is a value c of λ such that (a, b, c) is a critical point of F:

$$\text{(1)} \qquad \frac{\partial F}{\partial x}(a, b, c) = 0, \quad \frac{\partial F}{\partial y}(a, b, c) = 0, \quad \frac{\partial F}{\partial \lambda}(a, b, c) = 0.$$

Thus, in principle, one can find the candidates for the desired constrained extremum of f by solving the three simultaneous equations (1) for a, b, c. In the nicest situations there will be only one solution, which gives immediately the sought-for extremum (a, b) of f.

The aim here is to explain the geometric underpinning of the method. So assume $f(x, y)$ does have a maximum or a minimum at (a, b) under the constraint $g(x, y) = 0$. We shall assume further that (a, b) is a critical point of neither f nor g, the most common case. Note first that the partial derivatives of F are given by

$$\frac{\partial F}{\partial x} = \frac{\partial f}{\partial x} + \lambda \frac{\partial y}{\partial x}, \quad \frac{\partial F}{\partial y} = \frac{\partial f}{\partial y} + \lambda \frac{\partial g}{\partial y}, \quad \frac{\partial F}{\partial \lambda} = g.$$

The third equality in (1), therefore, just says that $g(a, b) = 0$, i.e., that (a, b) satisfies the constraint. The other two equalities in (1) can be written as

$$\text{(2)} \qquad \frac{\partial f}{\partial x}(a, b) = -c \frac{\partial g}{\partial x}(a, b), \quad \frac{\partial f}{\partial y}(a, b) = -c \frac{\partial g}{\partial y}(a, b).$$

What do these mean?

To shorten the notation, let's define

$$\alpha = \frac{\partial f}{\partial x}(a, b), \quad \beta = \frac{\partial f}{\partial y}(a, b), \quad \tilde{\alpha} = \frac{\partial g}{\partial x}(a, b), \quad \tilde{\beta} = \frac{\partial g}{\partial y}(a, b).$$

Rewritten in the new notation, (2) becomes

$$\text{(3)} \qquad \alpha = -c\tilde{\alpha}, \quad \beta = -c\tilde{\beta}.$$

Suppose for definiteness that (a, b) is a maximum of $f(x, y)$ under the constraint $g(x, y) = 0$, and let $m = f(a, b)$. Consider the level curve $f(x, y) = m$ (see Figure 3.1). It separates the region where f is larger than m from the region where f is smaller than m. On the level curve $g(x, y) = 0$ the function f takes no value larger than m, so that curve, although it touches the level curve $f(x,y)=m$ at (a,b), cannot pass through the latter curve; it must stay in the region where $f(x,y)\le m$. From this it follows that the two curves $f(x,y)=m$ and $g(x,y)=0$ share a common tangent line at the point (a,b) (see the figure).

The tangent lines at (a,b) to the curves $f(x,y)=m$ and $g(x,y)=0$ have the respective equations

$$\alpha(x-a)+\beta(y-b)=0, \quad \tilde{\alpha}(x-a)+\tilde{\beta}(y-b)=0 \tag{4}$$

(see Supplementary Notes 1). Now simple algebraic reasoning (left to the reader) shows that the two equations (4) define the same line if and only if the coefficients α,β are proportional to the coefficients $\tilde{\alpha},\tilde{\beta}$, i.e., there is a number γ such that $\alpha=\gamma\tilde{\alpha}$ and $\beta=\gamma\tilde{\beta}$. This gives (3) with $c=-\gamma$.

To summarize, the first two equalities in (1) just say that the level curves $f(x,y)=f(a,b)$ and $g(x,y)=0$ have a common tangent line at the point (a,b).