Lecture Notes for 3/18

He-Zhe Lin, March 6, 2021

Quadratic Approximation in Matrix Form

Recall: Linear approximation

Let $f(x, y)$ be a function differentiable at (x_0, y_0). The linear approximation of f at (x_0, y_0) is

$$\begin{aligned} f(x,y) &\approx f(x_0,y_0) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) \\ &= f(x_0,y_0) + \langle \nabla f(x_0,y_0), (x-x_0, y-y_0)\rangle, \end{aligned}$$

where $\langle \cdot, \cdot \rangle$ denotes the inner product of two vectors.

The meaning of linear approximation is that we may use linear functions to approximate a function near a given point, given the function is differentiable at the point. We now wonder whether we can approximate the function more accurately, given the function has more properties.

Quadratic Approximation

Assume $f(x, y)$ has continuous second-order partial derivatives at (x_0, y_0), i.e, f_{xx}, f_{xy}, f_{yx}, f_{yy} are continuous at (x_0, y_0). Then the function value of the points near (x_0, y_0) can be written as (we use $f_{xx}, f_{xy}, f_{yx}, f_{yy}$ to denote the second order partial derivatives at (x_0, y_0))

$$\begin{aligned} f(x,y) \approx & f(x,y) + f_x(x_0,y_0)(x-x_0) + f_y(x_0,y_0)(y-y_0) \\ & + \frac{f_{xx}}{2}(x-x_0)^2 + \frac{f_{xy}}{2}(x-x_0)(y-y_0) + \frac{f_{yx}}{2}(y-y_0)(x-x_0) + \frac{f_{yy}}{2}(y-y_0)^2 \end{aligned}$$

Let $u = \begin{bmatrix} x-x_0 \\ y-y_0 \end{bmatrix}$ and $H = \begin{bmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{bmatrix}$. Write the approximation above as matrix form, we have

$$f(x,y) \approx f(x_0,y_0) + \mathbf{u}^T \nabla f(x_0,y_0) + \frac{1}{2}\mathbf{u}^T H \mathbf{u}$$

Remark. The H matrix here is called the Hessian of the function.

Example Find the first- and second-degree Taylor polynomials L and Q of $f(x, y) = xe^y$ at $(0, 1)$. Compare the values of L, Q and f at $(0.9, 0.1)$.

Sol. We have that $f_x = e^y$, $f_y = xe^y$, $f_{xx} = 0$, $f_{xy} = f_{yx} = e^y$, $f_{yy} = xe^y$. Then

$$L(x,y) = f(1,0) + 1(x-1) + 1(y-0)$$

and

$$Q(x,y) = f(1,0) + 1(x-1) + 1(y-0) + \frac{1}{2}\left(0(x-1)^2 + 2(x-1)(y-0) + 1(y-0)^2\right)$$

Hence we have $f(0.9, 0.1) = 0.994653$, $L(0.9, 0.1) = 1$ and $Q(0.9, 0.1) = 1 + \frac{1}{2}(-0.01) = 0.995$.

Exercise Verify that the approximation has the same first- and second-order partial derivatives as f at (a, b).

Second Derivative Test Revisit

Theorem Suppose the second partial derivatives of f are continuous on a disk with center (a, b), and suppose that $f_x(a, b) = 0$ and $f_y(a, b) = 0$ (so (a, b) is a critical point of f). Let

$$D = D(a, b) = f_{xx}(a, b)f_{yy}(a, b) - [f_{xy}(a, b)]^2$$

(a) If $D > 0$ and $f_{xx}(a, b) > 0$, then $f(a, b)$ is a local minimum.
(b) If $D > 0$ and $f_{xx}(a, b) < 0$, then $f(a, b)$ is a local maximum.
(c) If $D < 0$, then (a, b) is a saddle point of f.
(d) If $D = 0$, then the test gives no information.

In fact, the criteria in (a) is equivalent to the positive definiteness of the Hessian of f at (a, b). We say the Hessian matrix is positive definite if for $v \in \mathbb{R}^2$

$$v^T H v > 0.$$

Similarly, we say that the Hessian matrix is negative definite if for $v \in \mathbb{R}^2$

$$v^T H v < 0$$

With this definition, we can use the quadratic approximation to view second derivative test:

$$f(x, y) \approx f(x_0, y_0) + \mathbf{u}^T \nabla f(x_0, y_0) + \frac{1}{2}\mathbf{u}^T H \mathbf{u}$$

Since the test requires $\nabla f(a, b) = 0$, the second term is always zero. If H is positive definite, the last term is always positive. In this case, the value of f near (a, b) is larger than the value of (a, b), i.e, (a, b) is a local minimum.
On the other hand, if H is negative definite, the last term is always non-positive. Hence f attains a maximum at (a, b).

Cases that Lagrange Multiplier doesn't apply

Exercise 14.8.34

Consider the problem of maximizing the function $f(x,y) = 2x + 3y$ subject to the constraint $\sqrt{x} + \sqrt{y} = 5$.

1. Try using Lagrange multipliers to solve the problem.

$$2 = \lambda \frac{1}{2\sqrt{x}}$$
$$3 = \lambda \frac{1}{2\sqrt{y}}$$
$$\sqrt{x} + \sqrt{y} = 5$$

   Solve the above equations, we obtain

$$\frac{\lambda}{4} + \frac{\lambda}{6} = \frac{5}{\lambda} = 5 \implies \lambda = 12,\ x = 9,\ y = 4$$

   $f(9,4) = 30$.

2. Does $f(25,0)$ give a larger value than the one in part (a)?
   Yes, since the gradient at $(25,0)$ is undefined.

Exercise 14.8.35

Consider the problem of minimizing the function $f(x,y) = x$ on the curve $y^2 + x^4 - x^3 = 0$.

1. Try using Lagrange multipliers to solve the problem.

$$1 = \lambda(4x^3 - 3x^2)$$
$$0 = \lambda(2y)$$
$$y^2 + x^4 - x^3 = 0$$

2. Show that the minimum value is $f(0,0) = 0$ but the Lagrange condition $\nabla f(0,0) = \lambda \nabla g(0,0)$ is not satisfied for any value of λ.

3. Explain why Lagrange multipliers fail to find the minimum value in this case.

Remark. Actually, the gradient of the restricted conditions need to be linearly independent.