Extrema of Two-Variable Functions and the Hessian Test

N-20260325

§1.1 Local extrema

Let $f : G \subseteq \mathbb{R}^2 \to \mathbb{R}$. A point $P_0 \in G$ is a **local maximum** if there is a neighborhood U of P_0 such that

$$f(P) \le f(P_0) \qquad (P \in U \cap G).$$

It is a **local minimum** if the inequality is reversed. If the inequality is strict for all nearby $P \neq P_0$, the extremum is strict.

The word "local" matters. A point can be best among nearby points without being best on the whole domain.

§1.2 First-order necessary condition

Theorem 1.2.1 (Fermat theorem in several variables)
If f is differentiable at an interior point $P_0 \in G$ and P_0 is a local extremum, then

$$\nabla f(P_0) = 0.$$

Proof. For every direction u, the one-variable function

$$\varphi(t) = f(P_0 + tu)$$

has a local extremum at $t = 0$. Hence $\varphi'(0) = 0$. But

$$\varphi'(0) = \nabla f(P_0) \cdot u.$$

This vanishes for every u, so $\nabla f(P_0) = 0$. □

A point with $\nabla f = 0$ is called a **critical point**. Critical points are candidates, not conclusions.

§1.3 Quadratic model and the Hessian

At a critical point P_0, the Taylor expansion begins with the quadratic term:

$$f(P_0 + h) = f(P_0) + \frac{1}{2} h^T H(P_0) h + o(\|h\|^2),$$

where

$$H(P_0) = \begin{pmatrix} f_{xx}(P_0) & f_{xy}(P_0) \\ f_{yx}(P_0) & f_{yy}(P_0) \end{pmatrix}.$$

When the second partial derivatives are continuous, $f_{xy} = f_{yx}$, so the Hessian is symmetric.

Thus the classification problem becomes a problem about a quadratic form.

§1.4 Second derivative test

Let

$$A = f_{xx}(P_0), \qquad B = f_{xy}(P_0), \qquad C = f_{yy}(P_0), \qquad D = AC - B^2.$$

Theorem 1.4.1 (Second derivative test)

Assume $\nabla f(P_0) = 0$ and the second partial derivatives are continuous near P_0. Then:

condition	conclusion
$D > 0,\ A > 0$	P_0 is a strict local minimum
$D > 0,\ A < 0$	P_0 is a strict local maximum
$D < 0$	P_0 is a saddle point
$D = 0$	the test is inconclusive

Proof skeleton. The quadratic part is

$$q(h,k) = Ah^2 + 2Bhk + Ck^2.$$

A symmetric 2×2 matrix is positive definite exactly when $A > 0$ and $D = AC - B^2 > 0$. It is negative definite exactly when $A < 0$ and $D > 0$. If $D < 0$, the quadratic form takes both positive and negative values. The Taylor expansion shows that the sign of the quadratic term controls the sign of $f(P_0 + h) - f(P_0)$ for sufficiently small h, except in the inconclusive case $D = 0$. □

Example 1.4.2

For

$$f(x,y) = x^2 + 3y^2,$$

the only critical point is $(0,0)$, and

$$H = \begin{pmatrix} 2 & 0 \\ 0 & 6 \end{pmatrix}.$$

Here $D = 12 > 0$ and $A = 2 > 0$, so $(0,0)$ is a strict local minimum.

Example 1.4.3

For

$$f(x,y) = x^2 - y^2,$$

the origin is critical, but

$$H = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}, \qquad D = -4 < 0.$$

The origin is a saddle point. Indeed, along the x-axis the function is positive, while along the y-axis it is negative.

Example 1.4.4 (Inconclusive case)

For $f(x, y) = x^4 + y^4$, the Hessian at the origin is the zero matrix, so $D = 0$. The test is inconclusive, but the origin is still a strict local minimum. For $g(x, y) = x^3 - y^3$, the Hessian at the origin is also zero, but the origin is not an extremum. This is why $D = 0$ really means "use another method."

§1.5 A picture of the Hessian test

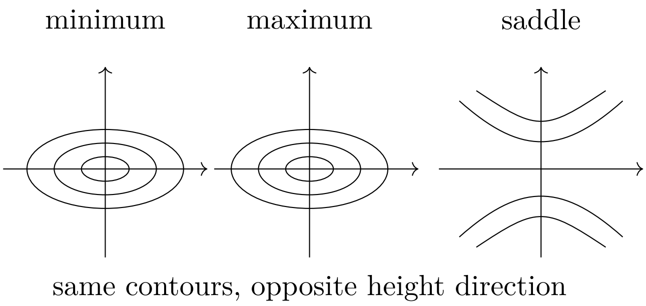

Figure 1.1: Near a nondegenerate critical point, the Hessian determines the local quadratic shape.

§1.6 Constrained extrema

Suppose one wants to optimize $f(x, y)$ subject to a constraint

$$g(x, y) = 0.$$

At a regular point of the constraint curve, $\nabla g \neq 0$, so ∇g is normal to the constraint. At a constrained extremum, moving along the constraint produces no first-order change in f. Thus ∇f must also be normal to the constraint.

Theorem 1.6.1 (Lagrange multiplier condition)

If P_0 is a constrained local extremum of f on $g = 0$, and $\nabla g(P_0) \neq 0$, then there exists λ such that

$$\nabla f(P_0) = \lambda \nabla g(P_0).$$

For several constraints $g_1 = \cdots = g_k = 0$, the gradient of f lies in the span of the constraint gradients:

$$\nabla f = \lambda_1 \nabla g_1 + \cdots + \lambda_k \nabla g_k.$$

Example 1.6.2

Find extrema of $f(x,y) = x + y$ on the circle $x^2 + y^2 = 1$. The equations are

$$(1,1) = \lambda(2x, 2y), \qquad x^2 + y^2 = 1.$$

Thus $x = y$, and $2x^2 = 1$, so

$$(x,y) = \left(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}\right) \quad \text{or} \quad (x,y) = \left(-\frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}\right).$$

The corresponding values are $\sqrt{2}$ and $-\sqrt{2}$, the maximum and minimum.

§1.7 Endpoint and boundary warnings

The equation $\nabla f = 0$ applies only to interior unconstrained extrema. The equation $\nabla f = \lambda \nabla g$ applies only at regular constrained points. If the domain has corners, endpoints, singular constraint points, or boundary pieces, they must be checked separately.

A reliable workflow is:

1. find interior critical points;
2. find constrained or boundary candidates;
3. check singular points and endpoints;
4. compare values when the problem asks for absolute extrema.