

12. UNCONSTRAINED OPTIMIZATION: SECOND-ORDER CONDITIONS AND THE HESSIAN

J.P. Bastos*

EPH 1100: Analytical Tools

Reading: Hoy et al., §11.5, §12.2.

- **§11.5:** The Hessian matrix, curvature, and definiteness conditions for concavity/convexity.
- **§12.2:** Second-order conditions for unconstrained optimization; determinant test for classifying critical points.

1. Why the FOC Is Not Enough

Setting both partial derivatives to zero gives a critical point, a candidate optimum, but not yet an answer to whether it is actually a maximum or a minimum.

Recall the single-variable case. If $f'(x^*) = 0$, we know x^* is a critical point. To determine whether it is a maximum or minimum, we check the second derivative: $f''(x^*) < 0$ means the function is concave at x^* (a hill top), and $f''(x^*) > 0$ means it is convex (a bowl bottom).

However, for functions of two variables, this is not enough. Consider the function $h(x, y) = x^2 - y^2$. The only critical point is $(0, 0)$. If we check curvature along the x -axis only (setting $y = 0$): $h(x, 0) = x^2$, a bowl. The origin looks like a *minimum*. If we check curvature along the y -axis only (setting $x = 0$): $h(0, y) = -y^2$, a hill. The origin looks like a *maximum*. Neither is globally correct. This is a case of a **saddle point** at the origin: it is a local minimum in the x -direction and a local maximum in the y -direction simultaneously. Checking h_{xx} alone (or h_{yy} alone) is not sufficient.

Thus for functions of two (or more) variables, the same logic applies but requires checking curvature in *every direction simultaneously*, not just along the coordinate axes.

2. The Hessian Matrix

The tool that captures curvature in all directions simultaneously is the Hessian matrix, which collects all the second-order partial derivatives.

*Assistant Professor of Economics, University of Austin (UATX). jbastos@uastin.org

Definition. The **Hessian matrix** of $f(x, y)$ at a point (x^*, y^*) is the 2×2 matrix of second-order partial derivatives:

$$H = \begin{pmatrix} f_{xx} & f_{xy} \\ f_{yx} & f_{yy} \end{pmatrix} = \begin{pmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial x \partial y} \\ \frac{\partial^2 f}{\partial y \partial x} & \frac{\partial^2 f}{\partial y^2} \end{pmatrix}$$

By Young's theorem, $f_{xy} = f_{yx}$, so the Hessian is symmetric.

For a function of n variables, the Hessian is an $n \times n$ symmetric matrix of all second-order partial derivatives $f_{ij} = \partial^2 f / \partial x_i \partial x_j$.

The connection to the single-variable case: the Hessian is the multivariable generalization of the second derivative. In one variable, we asked whether $f''(x^*) < 0$ or $f''(x^*) > 0$. In two variables, we ask whether the Hessian is **negative definite** (a hill, curving downward in every direction, a maximum) or **positive definite** (a bowl, curving upward in every direction, a minimum). The word “definite” is only emphasizing the literal meaning; it is “definitely” positive or negative in every direction.

3. The Determinant Test

For a 2×2 Hessian, definiteness is assessed via two conditions: the sign of the upper-left entry f_{xx} and the sign of the determinant $|H| = f_{xx}f_{yy} - (f_{xy})^2$.

Second-Order Conditions (2 variables). Suppose (x^*, y^*) is a critical point of $f(x, y)$. Let H be the Hessian evaluated at (x^*, y^*) , and let $|H| = f_{xx}f_{yy} - (f_{xy})^2$.

- **Local maximum:** $f_{xx} < 0$ and $|H| > 0$ (negative definite – a hill)
- **Local minimum:** $f_{xx} > 0$ and $|H| > 0$ (positive definite – a bowl)
- **Saddle point:** $|H| < 0$ (indefinite – curves both ways)
- **Test inconclusive:** $|H| = 0$

The intuition behind $|H|$ is similar to that of the single-variable case. A negative determinant already tells us we have a saddle point, with no need to check the sign of f_{xx} or f_{yy} at all. A positive determinant means f_{xx} and f_{yy} share the same sign – either both a hill or both a bowl – but does not say which; that is what the f_{xx} condition above is for. The cross partial f_{xy} enters through $(f_{xy})^2$: large cross-terms can “cancel” the diagonal curvature and turn what looks like a maximum (both $f_{xx} < 0$ and $f_{yy} < 0$) into a saddle, exactly the case worked out below.

However, $|H| > 0$ alone cannot distinguish a max from a min, because the same value of $|H|$ is compatible with both. Consider these two Hessians:

$$H_1 = \begin{pmatrix} -1 & 0 \\ 0 & -1 \end{pmatrix}, \quad |H_1| = (-1)(-1) - 0^2 = 1$$

$$H_2 = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}, \quad |H_2| = (1)(1) - 0^2 = 1$$

Both give $|H| = 1$, yet H_1 (both curvatures negative) is a maximum and H_2 (both curvatures positive) is a minimum. The determinant is identical in both cases, so we need the sign of f_{xx} to distinguish them. Technically, you can also check the sign of f_{yy} as well. Since we already know $|H| > 0$, we also know that both will have the same sign. If f_{xx} and f_{yy} had opposite signs, their product would be negative and $|H|$ would be negative too to begin with, i.e. the saddle case.

Example: Saddle Point. For $h(x, y) = x^2 - y^2$ at $(0, 0)$:

$$h_{xx} = 2, \quad h_{xy} = 0, \quad h_{yy} = -2 \quad |H| = (2)(-2) - 0^2 = -4 < 0$$

The determinant is negative: a saddle, as expected. Note h_{xx} and h_{yy} have opposite signs, exactly the “automatic saddle” case above.

Example: Two-Product Firm. The profit function $\pi(x, y) = 10x + 12y - x^2 - 2y^2 - xy$ has a unique critical point at $(x^*, y^*) = (4, 2)$. We confirm this is a maximum.

Step 1: Compute the second-order partial derivatives.

$$\pi_{xx} = -2, \quad \pi_{xy} = -1, \quad \pi_{yy} = -4$$

Step 2: Write the Hessian and its determinant.

$$H = \begin{pmatrix} -2 & -1 \\ -1 & -4 \end{pmatrix}, \quad |H| = (-2)(-4) - (-1)^2 = 8 - 1 = 7 > 0$$

Step 3: Check the sign conditions: $\pi_{xx} = -2 < 0$ and $|H| = 7 > 0$. The Hessian is negative definite: $(4, 2)$ is a **local maximum**. Since this is the only critical point and profit is very negative for large x or y (the quadratic terms dominate), it is also the global maximum.

At $(x^*, y^*) = (4, 2)$: the firm should produce 4 units of product 1, 2 units of product 2, and earns maximum profit of $\pi^* = 32$.

To see why *both* conditions were needed, suppose the cross-partial had instead been $\pi_{xy} = -3$ (with $\pi_{xx} = -2$ and $\pi_{yy} = -4$ unchanged):

$$|H| = (-2)(-4) - (-3)^2 = 8 - 9 = -1 < 0 \quad (\text{saddle point})$$

Both own partials are still negative – profit still looks like it is falling off in the x -direction alone and in the y -direction alone – but the cross-term is now strong enough that increasing x and y *together* raises profit. This is the economic content of the cross-partial: it captures whether the two choices interact so strongly that neither margin, taken alone, can be trusted.

4. Global vs. Local Optima

The determinant test identifies *local* optima, and in economics we usually care about *global* optima. The two coincide whenever the Hessian has the same definiteness everywhere, not just at the critical point. If H is negative definite everywhere, f is strictly concave, and its unique local maximum is also the unique global maximum. If H is positive definite everywhere, f is strictly convex, and its unique local minimum is also the unique global minimum.

When the Hessian's definiteness does not hold everywhere, there is no such shortcut, and every critical point has to be found, classified with the determinant test, and compared by function value. This is the same procedure as the single-variable case, where we compared every local maximum against the others and against the boundary values. We will avoid this complication at all costs, for two reasons.

Why is the Hessian almost never needed?

First, the functional forms economists use are chosen because their curvature is already known everywhere, before any critical point is found. For instance, a profit function built from concave revenue and convex costs is strictly concave throughout, so its critical point is automatically the global maximum. Similarly, a cost function is assumed strictly convex throughout, so its critical point is automatically the global minimum. In both cases the shape of the function settles the second-order question before a single second derivative is computed.

Second, most optimization problems in economics are *constrained*: consumers maximize utility subject to a budget, firms minimize costs subject to an output target, and so on. In these cases, the relevant second-order object is not this Hessian but the **bordered Hessian**, built from the Lagrangian. That test too is rarely computed explicitly, because it is normally settled by the shape of the objective function (quasi-concavity) rather than by a matrix calculation. Many textbooks gloss over the bordered Hessian, leaning on the economic intuition behind it instead – in practice, it is not a critical piece of the Lagrangian toolkit for economics. Long story short, the Hessian is best understood as the apparatus behind that shortcut, and not a computation you should expect to repeat problem set after problem set.¹

Practice Problems

1. For each function, find the critical points and classify each as a local max, local min, or saddle point.

(a) $f(x, y) = x^2 + y^2 - 4x - 6y + 13$

(b) $g(x, y) = -2x^2 - 3y^2 + 12x + 18y - 5$

(c) $h(x, y) = x^2 - 4xy + y^2$

¹Tangentially, a determinant test that extends to $n \geq 2$ variables becomes a sequence of n nested determinants (the “leading principal minors”) that must all satisfy a sign pattern. You can easily imagine the nightmare this is.

2. A monopolist sells in two markets. Profit is:

$$\pi(q_1, q_2) = (20 - q_1)q_1 + (15 - q_2)q_2 - (q_1 + q_2)^2$$

- (a) Find the critical point (q_1^*, q_2^*) .
 - (b) Use the determinant test to confirm it is a maximum.
 - (c) Compute π^* .
3. A firm chooses output q and quality s to maximize profit:

$$\pi(q, s) = 30q - 2q^2 + 8s - s^2 - qs$$

- (a) Find the critical point.
 - (b) Classify it using the SOC.
 - (c) Interpret the cross term $-qs$: what does it mean economically for quality to reduce the marginal return to output?
4. Without computing all derivatives, explain why the function $f(x, y) = x^4 + y^4$ cannot have a saddle point at the origin even though $f_{xx}(0, 0) = f_{yy}(0, 0) = 0$ (so the determinant test is inconclusive). What feature of this function guarantees the origin is a minimum?