*Res. Lett. Inf. Math. Sci.*, 2005, Vol. **8**, pp 189-196
Available online at http://iims.massey.ac.nz/research/letters/ 

# Hessian sufficiency for bordered Hessian

ERIC IKSOON IM

*Department of Economics, College of Business and Economics,*
*University of Hawaii at Hilo, USA*

eim@hawaii.edu

We show that the second–order condition for strict local extrema in both constrained and unconstrained optimization problems can be expressed solely in terms of principal minors of the (Lagrengean) Hessian. This approach unifies the determinantal tests in the sense that the second-order condition can be always given solely in terms of Hessian matrix.

## 1 Introduction

In the theory of constrained optimization, we use the bordered Hessian determinantal criterion to test whether an objective function has an extremun at a critical point. However, the required signs of the minors in the case of the constrained optimization are quite different from those in the case without constraints, which is somewhat confusing. In this paper, we show that when the constraints are twice differentiable we do not need the bordered Hessian at all for determinantal test. We only need the (Lagrangian) Hessian matrix for the determinantal test for both unconstrained and constrained optimization problems. This saves the unnecessary switching from the Hessian matrix to the bordered Hessian matrix for determinantal test for the second-order sufficient condition when the optimization problem is subject to constraints..

## 2 Discussion

To set the stage, first we formally state the standard constrained optimization problem and the second-order sufficient condition, then address the issue of unified sign requirements for the second-order condition for optimization.

Let $\phi : S \to \mathbb{R}$ be a real-valued function defined on a set $S$ in $\mathbb{R}^n$, and $g : S \to \mathbb{R}^m (m < n)$ a vector function defined on $S$. Let $c$ be an interior point of $S$ and let $\ell$ be a point in $\mathbb{R}^m$. Define the Lagrangian function $\psi : S \to \mathbb{R}$ by the equation

$$\psi(x) = \phi(x) - \ell' g(x) , \tag{1}$$

where $\phi(x)$ and $g(x)$ both are twice differentiable at *c*, the $m\times n$ Jacobian matrix $Dg(c)$ has a full row rank *m*. Then, the sufficient first- and second-order conditions for $\phi(x)$ to have a strict local minimum (maximum) at *c* are respectively (Magnus and Neudecker, pp.135-318 for example):

$$d\psi(c;dx) = Dg(c)dx = 0 \text{ and } g(c) = 0 \quad \text{for all } dx \in \Re^n, \tag{2}$$

$$d^2\psi(c;dx) = dx'H\psi(c)dx > 0 \text{ for all } dx \neq 0 \text{ satisfying } dg(c;dx) = 0. \tag{3}$$

where $H\psi \equiv H\phi - \sum_{i=m}^{m} \lambda_i Hg_i(c)$, and $D$ and $H$ are respective notations for Jacobian and Hessian operators.

The components of the second-order condition in (3) can be consolidated into one as follows. Partitioning $dx$ and $B_{m\times n} \equiv Dg(c)$ conformably as $dx = (dx'_m \vdots dx'_{n-m})'$ and $Dg(c) = (B_m \vdots B_{m,n-m})$, we can rewrite $dg(c) = 0$ in the second order condition:

$$dg(c,dx) = Dg(c)\,dx = B_m\,dx_m + B_{m,n-m}\,dx_{n-m} = 0$$

from which $dx_m = -B_m^{-1}B_{m,n-m}dx_{n-m}$. (4)

Incorporating (4) into (3), two components of the second-order condition are now consolidated into one:

$$d^2\psi(c,dx) = dx'_{n-m}Q'H\psi\,Q dx_{n-m} > 0 \ (<0) \qquad \text{for } dx \neq 0$$

where $Q = (-B'_{m,n-m}B_m^{-1} \vdots I_{n-m})'$, or, equivalently $\Omega(c) \equiv Q'H\psi(c)Q > 0 \quad (<0)$. (5)

The strict inequality holds if and only if the signs of the principal minors of $\Omega(c)$ adhere to:

$|\Omega_k| > 0$ for strict local minima;

$(-1)^k|\Omega_k| > 0$ for strict maxima ( $k$ = m+1, $m$+2, …, $n$-$m$ ) (6)

where $\Omega_k = E_k\Omega(c)E'_k$; in which $E_k = \left(I_k \vdots 0_{k\times(n-k)}\right)$ and $Q_k = QE'_k$.

In practice, however, we test for the sign definiteness of the second-order condition, using the determinantal criterion based on the bordered Hessian matrix:

$(-1)^m\left|\overline{H}_r\psi(c)\right| > 0$ for strict local minima;

$(-1)^r\left|\overline{H}_r\psi(c)\right| > 0$ for strict local maxima, $(r = m+1,\ldots,n)$ (7)

$$\text{where: } \overline{H}_r\psi(c) = \begin{bmatrix} 0_{m\times m} & B_{m\times r} \\ B'_{m\times r} & \left(H\psi(c)\right)_r \end{bmatrix} \tag{8}$$

defines the principal submatrices of bordered Hessian of order $m+r$ of the bordered Hessian matrix of order $m+n$:

$$\overline{H}\psi(c)=\begin{bmatrix}0_{m\times m} & B_{m\times n}\\ B'_{m\times n} & H\psi(c)\end{bmatrix} \tag{9}$$

in which $B_{m\times r}$ denotes the first $r$ columns of $B_{m\times n}\equiv D\,g(c)$

However, the sufficient second-order condition for a strict local optimum can be stated purely in terms of principal minors of $H\psi(c)$ instead of those of the bordered Hessian as discussed in the following section.

## 3 Hessian Sufficiency for Bordered Hessian

In the Hessian alternative to the bordered-Hessian, it is essential to note that there is a rank condition implicit in the first-order condition, which is not needed in the bordered Hessian approach. To make the point, re-express the first part of first-order condition in (2) in variable form for further differentiation:

$$d\psi(x;dx)=D\psi(x)\,dx=0 \quad . \qquad \text{for all } dx_{n-m}\in\mathbb{R}^{n-m}$$

Differential of this first-order condition at $x=c$, with (4) incorporated, can be written as:

$$\begin{aligned} d^2\psi(c,dx) &= dx'\,H\psi(c)\,dx\\ &= dx'_{n-m}\Omega(c)\;dx_{n-m}=0 \qquad \text{for all } dx_{n-m}\in\Re^{n-m} \end{aligned} \tag{10}$$ [†]

Though Equation (10) contains the Hessian matrix, it is not the second-order condition. It is in essence the first-order condition in quadratic form which facilitates identifying a useful property of $Q'\,H\psi(c)\,Q$ associated with a strict local critical point.

Suppose that $dx_{n-m}\neq 0$. Then, $\Omega(c)\,dx_{n-m}=0 \qquad \text{for all } dx_{n-m}\neq 0$ (11)

Since $dx_{n-m}\neq 0$, $\Omega$ must be singular, hence $\Omega=\Omega(c)$ (therefore $c$) depends on $dx_{n-m}$ which violates the assumption that $c$ is a strict local critical point. Therefore, $\Omega$ should be nonsingular, which in turn implies $dx_{n-m}=0$. Hence, a lemma can be stated without proof:

**LEMMA**: Let $c$ be a strict local critical point for the optimization problem (1). Then, $\Omega(c)\equiv Q'\,H\psi(c)\,Q$ is nonsingular.

---

[†] Note that if $g(x)$ is linear, $H\psi$ does not depend on $x$, therefore $c$. This is exactly the reason why $g(x)$ is assumed to be twice differentiable. However, a linear constraint is technically twice differentiable when squared. Hence, our approach applies to both linear and nonlinear constraints.

Having established formally that $\Omega(c)$ is nonsingular if $c$ is a strict local critical point, we are now ready to restate the second-order sufficient condition in (5) solely in term of the (Lagrangian) Hessian instead of bordered Hessian as follows.

**THEOREM (General Sufficient Condition)**: If $H\psi(c)$ *is* nonnegative definite (nonpositive definite) where $c$ is a strict critical point $x$, $\psi(c)$ is a strict minimum (maximum), and $Rank(H\psi(c)) \geq n-m$

**Proof :** Since $c$ is a strict critical point, $\Omega(c) = Q' H\psi(c) Q$ is nonsingular (therefore of full rank) in view of the Lemma. If $H\psi(c)$ is nonnegative definite, it has a full-rank factorization: $H\psi(c) = TT'$ where $T$ is of full column rank. Hence, we can rewrite $\Omega(c) = (T'Q)'(T'Q)$ which is nonnegative definite. Since $\Omega(c)$ is both non-singular and nonnegative definite, $\Omega(c) > 0$ which implies that $\psi(c)$ is a strict local maximum. The proof for a strict local maximum is similar. The second part of the theorem can be proved as follows. $\Omega(c) > 0$ implies $Rank(H\psi(c)) \geq n-m$: Since $Rank(\Omega(c)) = n-m$, $Rank(T'Q) = n-m$. However, since $Rank(Q) = n-m$ by assumption, $Rank(T) \geq n-m$ from which follows $Rank(H\psi(c)) \geq n-m$. )

*QED*

**COROLLARY (Special Sufficient Condition):** If $H\psi(c)$ *is* positive definite (negative definite) where $c$ is a strict local critical point, $\psi(c)$ is a strict minimum (maximum), and $Rank(H\psi(c)) = n$.[‡]

**Proof**: Since positive definiteness (negative definiteness) is a special case of nonnegativity (nonpositivity), the corollary follows from the theorem as a trivial case. Since definite matrices are of full rank, $Rank(H\psi(c)) = n$. *QED*

## 4 An Example

Solve the problem (Magnus & Neudecker, pp. 138-139)

$$\underset{x=(v,w)}{Max(Min)} \quad \phi(x) = v^2 + w^2$$

subject to $v^2 + vw + w^2 = 3$.

Then, the Lagrangian function will be: $\psi(x) = v^2 + w^2 - \ell(v^2 + vw + w^2 - 3)$,

[‡]The special sufficient condition has been known and noted in a number of places (e.g., Hal R. Varian (p. 498)), but not the general sufficient condition.

and the first-order conditions are:

$$2v - 2\ell v - w\ell = 0; \quad 2w - \ell v - 2\ell w = 0; \quad v^2 + vw + w^2 = 3 .$$

Solving the first-order conditions for control variable vector $x = (v, w)$, we have three strict local stationary points: $x = c_1, c_2, c_3$ are defined respectively as

$$c_1 = (v_1, w_1) = (-1, -1), \; c_2 = (v_2, w_2) = (\sqrt{3}, -\sqrt{3}), \text{ and } c_3 = (v_3, w_3) = (-\sqrt{3}, \sqrt{3}) .$$

The corresponding Lagrangian multipliers are, respectively, $\ell_1 = 2/3$, $\ell_2 = 2$, and $\ell_3 = 2$. The (Lagrangian) Hessian and bordered Hessian are, respectively:

$$H\psi(x) = \begin{bmatrix} 2-2\lambda & -\lambda \\ -\lambda & 2-2\lambda \end{bmatrix}; \quad \overline{H}\psi(x) = \begin{bmatrix} 0 & 2v+w & v+2w \\ 2v+w & 2-2\lambda & -\lambda \\ v+2w & -\lambda & 2-2\lambda \end{bmatrix}.$$

Now, we check the sign behavior of all minors, not just the leading minors, of $H\psi(x)$ at each of the three critical points to see whether we have a local minimum or maximum or neither in view of the Lemma.

Minors of $H\psi(c_1)$ for the first stationary point are:

$$\left|\tilde{M}_1\right| = 2 - 2(2/3) = 2/3 > 0; \quad \left|\tilde{M}_2\right| = 0 .$$

$H\psi(c_1)$ is positive semi-definite of rank 1, hence, in view of the Theorem, the first critical point is a strict local minimum.

Corresponding principal minors of $H\psi(c_2)$ and $H\psi(c_3)$ are identical:

$$\left|\tilde{M}_1\right| = -2 < 0; \quad \left|\tilde{M}_2\right| = 0 .$$

Both $H\psi(c_2)$ and $H\psi(c_3)$ are negative semi-definite of rank 1, thus in light of the Theorem $\phi(x)$ has two strict local maxima.

(The same conclusion is arrived for each critical point by way of the bordered determinantal criterion in Magnus and Neudecker).

## 5 Linear Constraints as Twice Differentiables

Typical linear constraints can be transformed into twice differentiables by squaring the linear constraints. Therefore, the Hessian alternative to bordered Hessian applies to all constraints in general. This can be illustrated by a simple example in Chiang and Wainwright (Example 2, p. 360):

Find the extremum of

$$z = vw \text{ subject to } v + w = 6. \quad \text{(I)}$$

The Hessian and bordered Hessian matrices are respectively:

$$H\psi(x) = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}; \quad \overline{H}\psi(x) = \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}.$$

The determinant of bordered Hessian is positive: $\left|\overline{H}\psi(x)\right| = 2 > 0$ which meets the sign requirement for a strict local maximum, whereas the leading principal minors of Hessian are: $\left|H_1\psi(x)\right| = 0;$ $\left|H_2\psi(x)\right| = -1 < 0$ which do not meet the sign requirements for a local maximum. Clearly, the Hessian approach fails to identify the local maximum. The reason is that theorem is applicable only when the constraint is twice differentiable.

This problem can reformulated using the squared constraint as

$$z = vw \text{ subject to } (v + w)^2 = 6^2 \text{ or } v^2 + 2vw + w^2 = 36. \quad \text{(II)}$$

Then, the Lagrangian function for (II) will be: $\psi(x) = vw - \ell(v^2 + 2vw + w^2 - 36)$, and the first-order conditions are:

$$w - 2v\ell - 2w\ell = 0; \quad v - 2\ell v - 2\ell w = 0; \quad v^2 + 2vw + w^2 = 36.$$

The Lagrangian Hessian then will be:

$$H\psi(x) = \begin{bmatrix} -2\ell & 1-2\ell \\ 1-2\ell & -2\ell \end{bmatrix}; \quad \overline{H}\psi(x) = \begin{bmatrix} 0 & -2(v+w) & -2(v+w) \\ -2(v+w) & -2\ell & 1-2\ell \\ -2(v+w) & 1-2\ell & -2\ell \end{bmatrix}.$$

Solving the first-order conditions for control vector $x = (v, w)$, we have two strict local critical points:

$c_1 = (v, w) = (3, 3)$ and $c_2 = (v, w) = (-3, -3)$ and $\ell = 1/4$ for both critical points.

However, since $c_2$ does not satisfy the original constraint, it will be out of consideration. At $c_1 = (v, w) = (3,3)$,

$$H\psi(c_1) = \begin{bmatrix} -2\ell & 1-2\ell \\ 1-2\ell & -2\ell \end{bmatrix} = \begin{bmatrix} -1/2 & 1/2 \\ 1/2 & -1/2 \end{bmatrix}$$

for which: $\left|\tilde{M}_1\right| = -1/2 < 0; \quad \left|\tilde{M}_2\right| = 0$

$H\psi(c)$ is negative semi-definite of rank 1, hence, in view of the Theorem, the first critical point is a strict local maximum, which is exactly the same results as would be produced if we used the bordered Hessian determinantal rule to Problem (II).

The Hessian Approach to Problem I breaks down for the simple reason that the constraint is not twice differentiable as required by the Theorem, but it works for Problem II since the constraint are converted into a twice differentiable.

## 6 Conclusion

In this paper, we have shown that the conventional second-order sufficient condition for constrained optimization in terms of minors of bordered Hessian can be recast alternatively in terms of minors of Hessian matrix, making it possible to represent the second-order condition in general all in terms of minors of the Hessian matrix. This theoretical result dispels the misconception that borders in the bordered Hessian matrix have some bearings on the second-order condition. Furthermore, the second-order condition based on Hessian matrix is more convenient in practice than the bordered Hessian since finding minors of Hessian is less cumbersome than those of bordered Hessian. Practitioners of the determinantal tests in optimization theory find it a bit confusing to use the minors of Hessian for constrained optimization problems but minors of bordered Hessian for unconstrained problems. Finally, the theorem seemingly applicable only to the optimization problems with twice differentiable constraints indeed applies to problems with linear constraints, therefore applicable in general.

**APPENDIX**

As shown in Magnus and Neudecker ( p.55),

$$(-1)^m \left|\overline{H}_r\psi(x)\right| = (-1)^m \left|T_k\right|^2 \left|\Omega_k\right| \qquad (r = m + k;\ k = 1, 2, ..., n - m) \tag{A1}$$

where $T_k$ (nonsingular) is defined as: $T_k = \begin{pmatrix} B_{m\times m} & B_{m\times k} \\ 0_{k\times m} & I_k \end{pmatrix}$. (A2)

From (A1) for $k = n - m$ readily follows: $(-1)^m \left|\overline{H}_n \psi(x)\right| = \left|T_{n-m}\right|^2 \left|\Omega_{n-m}\right|$. (A.3)

It is clear from (A.3) that $|\Omega| \neq 0$ if and only if $\left|\overline{H}\psi(x)\right| \neq 0$, or equivalently $\Omega$ is nonsingular if and only if $\overline{H}\psi(x)$ is nonsingular. Note that $\overline{H}\psi = \overline{H}_n\psi(x)$ and $\Omega = \Omega_{n-m}$ which is obvious from the context.

## References

Chiang A. and K. Wainwright (2005). *Fundamental Methods of Mathematical Econmics*. Fourth Edition. McGraw-Hill, New York.

Magnus, J. R. and H. Neudecker (1998). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. Revised Edition. Wiley, New York.

Varian, H. R. (2001). *Microeconomic Analysis*. Norton & Company, New York.