Economics Letters 121 (2013) 440–443

Contents lists available at ScienceDirect

Economics Letters

journal homepage: www.elsevier.com/locate/ecolet

On second order conditions for equality constrained extremum problems

David M. Mandy *

Department of Economics, University of Missouri, 118 Professional Building, Columbia, MO 65203, USA

HIGHLIGHTS

- Establishes a simple relationship between a Hessian and bordered Hessian.
- Derives necessary and sufficient second order conditions from this relationship.
- The only proof that avoids use of quadratic forms subject to side conditions.
- Clarifies Samuelson's "great loss of symmetry".

ARTICLE INFO

Article history:
Received 4 July 2013
Received in revised form
2 September 2013
Accepted 13 September 2013
Available online 20 September 2013

JEL classification:
C60

Keywords:
Optimization
Hessian
Minors

ABSTRACT

We prove a relationship between the bordered Hessian in an equality constrained extremum problem and the Hessian of the equivalent lower-dimension unconstrained problem. This relationship can be used to derive principal minor conditions for the former from the relatively simple and accessible conditions for the latter.

It is surprisingly difficult to find a complete and accessible proof of conditions on principal minors of the bordered Hessian matrix for extremum problems with an arbitrary number n of choice variables and $m < n$ equality side conditions. Most presentations emphasize sufficiency and stop short of proving principal minor conditions, typically proving that definiteness of the matrix subject to side conditions is part of the sufficient conditions for an extremum but providing only an unproven statement that such definiteness is equivalent to particular sign patterns among principal minors. Lancaster (1968) is an exception, providing proofs of both the former (Section 4.5) and latter (Section R6.3), but Lancaster does not prove necessary second order conditions. Simon and Blume provide a relatively complete, accessible and modern treatment of both necessary and sufficient (semi) definiteness conditions for an unconstrained optimum (Simon and Blume, 1994, Section 30.4); and prove the equivalent sufficient principal minor conditions (pp. 393–395) but state without proof the equivalent necessary principal minor conditions (p. 383); and for the constrained case provide a full proof only that a definite Hessian subject to side constraints is part of the sufficient conditions (Section 30.5). One must typically supplement this type of discussion with proofs from Debreu (1952) or Gantmacher (1959, pp. 306–307) of the relationships between principal minors and (semi) definiteness subject to side conditions.

The traditional emphasis on sufficiency rather than necessity is both dated and incomplete for economic analysis.

It is dated because analytic derivation of the bordered Hessian may not be possible for a modern large scale optimization or equilibrium application, and even if the matrix can be derived checking definiteness is computationally complex (Pardalos and Schnitger, 1988). Computationally and analytically simpler conditions are available for applied problems (Magnus and Neudecker, 1988, pp. 135–139, Morrow, 2011).

It is incomplete because sufficiency is not useful for derivation of refutable hypotheses in optimization-based positive economic theories. These theories begin with a behavioral postulate that an economic agent engages in constrained optimizing behavior and then seek to uncover the refutable hypotheses implied by that

* Tel.: +1 573 882 1763.
E-mail address: mandyd@missouri.edu.

0165-1765/$ – see front matter © 2013 Elsevier B.V. All rights reserved.
http://dx.doi.org/10.1016/j.econlet.2013.09.017

behavioral postulate. Only necessary, not sufficient conditions can be implied by the optimization postulate. Principal minors of the bordered Hessian appear in many comparative static expressions, hence we must know necessary conditions on those principal minors to determine which comparative static signs are implied by the postulated optimizing behavior. This distinction between the role of necessary second order conditions in positive economic theory versus the role of sufficient conditions in computational applications is almost absent from the literature.

These comments apply with equal force to second order conditions for unconstrained extrema. But it is, of course, easier to prove (semi) definiteness of the Hessian for an unconstrained problem and also easier to prove the relationships between such (semi) definiteness and signs of principal minors. Hence one way to assemble a relatively accessible and complete proof of principal minor conditions when there are equality constraints, both necessary and sufficient, is to apply the corresponding conditions for the unconstrained case to the lower-dimension version of the constrained problem obtained by substituting the constraint for an equal number of the choice variables. This approach suffers, however, from the relative disutility of the resulting conditions. Generally, conditions on the bordered Hessian are easier to apply and more immediately usable, as originally emphasized by Hancock (1917, Chapter 6) and Samuelson (1947, Appendix A), because it is relatively easy to obtain comparative static expressions by differentiating Lagrangian-based first order conditions, and it is the principal minors of the bordered Hessian that appear directly in those derivations.

This paper derives principal minor conditions on the bordered Hessian directly from the corresponding conditions on the lower-dimension unconstrained problem. It builds naturally on textbook presentations of the unconstrained case, supplemented by the unconstrained relationships between (semi) definiteness and principal minors, without using any concepts except calculus and matrix algebra familiar to graduate students in economics. In particular, the approach completely eliminates consideration of quadratic forms subject to side conditions. Such consideration is unavoidable with extant proofs but is tedious and repetitive once principal minor conditions have been established for (semi) definite quadratic forms without side conditions. The new proof integrates constrained and unconstrained statements of principal minor conditions, both necessary and sufficient. It is similar in spirit to Im (2005) but Im addresses only the sufficiency relationship between the Hessian and bordered Hessian and does not explore the relationships between the principal minors of the two matrices.

1. Notation

Denote by A' the transpose of the matrix A. If A is square of dimension k, $A^{(J)}$ denotes the principal submatrix consisting of rows and columns $J \subset \{1, \dots, k\}$ (excluding $J = \emptyset$).

Assume $X \subset \mathbb{R}^n$ is an open set via the usual Euclidean metric; let $f : X \to \mathbb{R}^1$ be the objective and $h : X \to \mathbb{R}^m$ for $m < n$ be the equality constraint. The values of a vector-valued function like h are interpreted as column vectors in matrix equations. $\mathbf{D}h(x)$ is the $m \times n$ Jacobian matrix of h evaluated at $x \in X$ (Apostol, 1974, p. 351). To avoid notational clutter, the arguments of a Jacobian are omitted when the point of evaluation is clear from the context. The Lagrangian function $L : X \times \mathbb{R}^m \to \mathbb{R}^1$ is defined by $L(x, \lambda) \equiv f(x) - \lambda' h(x)$.

Let $x = (y, z)$ be a partition of the choice vector x into the first $n - m$ and last m components so the constraint can be substituted for z to form the lower-dimension unconstrained problem. When differentiating with respect to part of a partition, notation like $\mathbf{D}_y h(y, z)$ denotes the $m \times (n - m)$ Jacobian of h with respect to y evaluated at (y, z). A critical point is denoted $\hat{x} = (\hat{y}, \hat{z}) \in X$.

For a real-valued function like f, the $n \times n$ matrix $\mathbf{D}^2 f(x) \equiv \mathbf{D}(\mathbf{D}f)'(x)$ is the Hessian matrix of f evaluated at $x \in X$. As above, $\mathbf{D}_y^2 f(y, z)$ denotes the $(n-m) \times (n-m)$ Hessian of f with respect to y evaluated at (y, z). Recall that the Hessian is symmetric at x when it both exists in a neighborhood of x and each entry is continuous at x (Apostol, 1974, Theorem 12.13).

2. Lower-dimension objective

Restating an equality constrained optimization problem as a lower-dimension unconstrained objective with an optimum that can be studied using calculus requires that it be possible to substitute the constraint for an equal number of choice variables while retaining differentiability within a neighborhood of the optimum. A key part of the sufficient conditions for this from the implicit function theorem is that $\mathbf{D}h(\hat{x})$ has rank m (Apostol, 1974, Theorem 13.7). Put in a more convenient form, it must be possible to order the choice variables so that the partition $x = (y, z)$ yields a nonsingular $\mathbf{D}_z h(\hat{y}, \hat{z})$ matrix.

This rank condition may appear to be an aspect of the reduced-dimension approach that is not needed in the Lagrangian approach, or it may appear that substitution of h for z introduces a loss of symmetry in the treatment of choice variables.[1] However, this rank condition is exactly the constraint qualification typically used to ensure existence of a unique Lagrange multiplier; without it, manipulations involving Lagrange multipliers are ill-defined (see, for example, Apostol, 1974, Theorem 13.12). And the classical bordered Hessian approach is equally "unsymmetric" when properly interpreted, precisely because the classical conditions are not applied to the portion z of the choice vector (more on this at the end of Section 4).

Assuming h is continuously differentiable in an open ball about $\hat{x}$ and $\mathbf{D}_z h(\hat{y}, \hat{z})$ is nonsingular, the implicit function theorem ensures the existence of an open ball $B_\epsilon(\hat{y}) \subset \mathbb{R}^{n-m}$ and a unique continuously differentiable function $\phi : B_\epsilon(\hat{y}) \to \mathbb{R}^m$ such that $(\hat{y}, \phi(\hat{y})) = \hat{x}$ and

$$h(y, \phi(y)) = 0 \quad \text{for } y \in B_\epsilon(\hat{y}). \tag{1}$$

Therefore $\hat{y}$ is a local extremum of the composite function $\tilde{f}(y) = f(y, \phi(y))$ over $y \in B_\epsilon(\hat{y})$ if and only if $\hat{x}$ is a local extremum of f subject to $h = 0$. $\tilde{f}$ is the lower-dimension unconstrained objective.

3. Relationship between Hessian and Bordered Hessian

The following theorem provides the algebraic relationship between the Hessian of $\tilde{f}$ and the Hessian of L at a stationary point of L.

Theorem 1. *For $\hat{x} = (\hat{y}, \hat{z}) \in X$, assume:*

1. *$h(\hat{x}) = 0$,*
2. *$\mathbf{D}^2 f$ and $\mathbf{D}^2 h$ exist in an open ball about $\hat{x}$ and are continuous at $\hat{x}$, and*
3. *$\mathbf{D}_z h(\hat{y}, \hat{z})$ is nonsingular.*

If there exists $\hat{\lambda} \in \mathbb{R}^m$ such that $\mathbf{D}_z L(\hat{y}, \hat{z}, \hat{\lambda}) = 0$ then

$$\mathbf{D}^2 \tilde{f} = \mathbf{D}_y^2 L - (\mathbf{D}_y(\mathbf{D}_{(z,\lambda)} L)')' (\mathbf{D}_{(z,\lambda)}^2 L)^{-1} \mathbf{D}_y(\mathbf{D}_{(z,\lambda)} L)' \quad \text{at } (\hat{x}, \hat{\lambda}). \tag{2}$$

[1] Hancock criticized the lower-dimension approach as "unsymmetric" in the treatment of choice variables (Hancock, 1917, p. 103). Samuelson, who apparently relied on Hancock in this regard, declares: "However, there is a great loss of symmetry in such a procedure since not all our variables are treated alike. Fortunately, by the use of an artifice which can be rigorously justified, it is possible to derive a more symmetrical set of conditions" (Samuelson, 1947, p. 363) (the "artifice" referred to here is the Lagrangian function).