Advanced Micro: Recitation 4
Second Order Conditions

Motaz Al-Chanati

February 22, 2019

1 Quadratic Forms

1.1 Definition

A **quadratic form** in two variables is a function of the form:

$$Q(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2$$

Note we can write this in matrix form:

$$Q(x_1, x_2) = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = x'Ax$$

We can take A to be symmetric without loss of generality - if we had two values b_1 and b_2, we could replace them both with $\frac{1}{2}(b_1 + b_2)$ without changing the value of Q.

We can generalize this to n dimensions. A quadratic form on $\mathbb{R}^n$ is a function of the form:

$$Q(x) = \sum_{i=1}^{n} \sum_{j=1}^{n} a_{ij} x_i x_j = x'Ax$$

where A is a symmetric $n \times n$ matrix.

1.2 Definiteness

One important property a quadratic form can have is that of definiteness. We say a quadratic form $Q(x) = x'Ax$ (or simply that the matrix A) is:

- **positive definite** if $Q(x) > 0$ for all $x \neq 0$
- **positive semi-definite** if $Q(x) \geq 0$ for all $x \neq 0$
- **negative definite** if $Q(x) < 0$ for all $x \neq 0$

- **negative semi-definite** if $Q(x) \leq 0$ for all $x \neq 0$
- **indefinite** if none of the above hold

Let's consider the 2×2 case. Suppose $a \neq 0$. Then we have (by completing the square):

$$\begin{aligned} Q(x) &= ax_1^2 + 2bx_1x_2 + cx_2^2 \\ &= a\left(x_1^2 + 2\frac{b}{a}x_1x_2 + \frac{c}{a}x_2^2\right) \\ &= a\left[\left(x_1 + \frac{b}{a}x_2\right)^2 + \frac{ac-b^2}{a^2}x_2^2\right] \end{aligned}$$

The only terms whose signs we don't know are a and $ac-b^2$ (everything else in the expression is non-negative). If $a > 0$ and $ac - b^2 = \det(A) > 0$, then $Q(x) > 0$ for all $x \neq 0$, so Q is positive definite. Similarly, if Q is positive definite, setting $x_2 = 0$ implies $a > 0$. Setting $x_1 = -(b/a)x_2$ implies that $ac - b^2 > 0$ (since we know that $a > 0$).

- $Q(x)$ is positive definite $\Leftrightarrow a > 0$ and $\det(A) > 0$
- $Q(x)$ is negative definite $\Leftrightarrow a < 0$ and $\det(A) > 0$

What about positive semi-definite? A similar argument shows:

- $Q(x)$ is positive semi-definite $\Leftrightarrow a \geq 0, c \geq 0$ and $\det(A) \geq 0$
- $Q(x)$ is negative semi-definite $\Leftrightarrow a \leq 0, c \leq 0$ and $\det(A) \geq 0$

Example. Determine the definiteness of $A = \begin{pmatrix} 2 & 1 \\ 1 & 4 \end{pmatrix}$.

We see $2 > 0$ and $\det(A) = 8 - 1 = 7 > 0$, so A is positive definite.

Example. Show that for any positive (negative) definite matrix, the diagonals must all be positive (negative).

Take a matrix A where a_{ii} is its i^{th} diagonal element. Since $Q(x)$ has to apply for any $x \neq 0$, it must also apply to the vector $x = e_i$ (a vector with 1 in the i^{th} element and 0 elsewhere). We have that $Q(e_i) = e_i'Ae_i = a_{ii}$. Therefore, for a positive definite matrix, we must have that $Q(e_i) = a_{ii} > 0$ and for a negative definite matrix $a_{ii} < 0$. Since this was for any arbitrary i, then this must be true for any element on the diagonal.

1.3 Testing Definiteness

There is a general test for the definiteness of any quadratic form, but we first need to introduce two concepts.

Principal Minor A principal minor of order k of an $n \times n$ matrix A is the *determinant* of the $k \times k$ submatrix consisting of deleting $n-k$ rows and the corresponding $n-k$ columns of A (or equivalently, of keeping k rows and the corresponding k columns of A).

Leading Principal Minor A leading principal minor of order k of an $n \times n$ matrix A is the *determinant* of the $k \times k$ sub-matrix consisting of the first k rows and columns of A.

Example. The matrix $A = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix}$ has three principal minors of order 2:

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21} \qquad \text{(delete row 3, column 3)}$$

$$\begin{vmatrix} a_{11} & a_{13} \\ a_{31} & a_{33} \end{vmatrix} = a_{11}a_{33} - a_{13}a_{31} \qquad \text{(delete row 2, column 2)}$$

$$\begin{vmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{vmatrix} = a_{22}a_{33} - a_{23}a_{32} \qquad \text{(delete row 1, column 1)}$$

Of these, the first one is the leading principal minor of order 2.

With this in hand, we will state a result about the definiteness of matrices:

Theorem 1. *Let A be a symmetric matrix and let $Q(x) = x'Ax$ be a quadratic form. Let D_k be the leading principal minor of order k, and let Δ_k denote an arbitrary principal minor of order k. Then:*

- *Q is positive definite $\Leftrightarrow$ $D_k > 0$ for $k = 1, ..., n$*
- *Q is positive semi-definite $\Leftrightarrow$ $\Delta_k \geq 0$ for all principal minors of order $k = 1, ..., n$*
- *Q is negative definite $\Leftrightarrow$ $(-1)^k D_k > 0$ for $k = 1, ..., n$*
- *Q is negative semi-definite $\Leftrightarrow$ $(-1)^k \Delta_k \geq 0$ for all principal minors of order $k = 1, ..., n$*

A few notes:

- This is the exactly generalization of the 2-dimensional case we gave earlier (check!)
- To check for positive semi-definiteness, it is not enough to look at the leading principal minors! You need to check <u>all</u> the principal minors. Note for the 2×2 case we had a condition on a and c: these are the two principal minors of order 1.
- The condition for negative definite is just saying that the leading principal minors alternate in sign (starting with negative for order 1). For negative semi-definite, it's the same idea: every principal minor of odd orders is non-positive, and of every principal minor of even orders is non-negative.

1.4 Bordered Matrices

Another way to look at definiteness is as follows:

- $x = 0$ is the global minimum of Q if and only if Q is positive definite
- $x = 0$ is the global maximum of Q if and only if Q is negative definite

However, the idea here works when x is allowed to be anything in $\mathbb{R}^n$. Often, we want to look at maximizing or minimizing a function subject to a constraint. For this, we will introduce the idea of a bordered matrix.

Let's consider the simple two dimensional example:

$$Q(x_1, x_2) = ax_1^2 + 2bx_1x_2 + cx_2^2 = \begin{pmatrix} x_1 & x_2 \end{pmatrix} \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$$

Constrained to the linear subspace: (where α and β are constants)

$$\alpha x_1 + \beta x_2 = 0$$

We could re-arrange this to get $x_1 = -\frac{\beta}{\alpha} x_2$ and plug it into Q:

$$\begin{aligned} Q\left(-\frac{\beta}{\alpha}x_2, x_2\right) &= a\left(-\frac{\beta}{\alpha}x_2\right)^2 + 2b\left(-\frac{\beta}{\alpha}x_2\right)x_2 + cx_2^2 \\ &= \left(\frac{a\beta^2}{\alpha^2} - 2b\frac{\beta}{\alpha} + c\right)x_2^2 \\ &= \frac{a\beta^2 - 2b\alpha\beta + c\alpha^2}{\alpha^2}x_2^2 \end{aligned}$$

This tells us that Q is positive definite on the constraint set if and only if the numerator $a\beta^2 - 2b\alpha\beta + c\alpha^2 > 0$ and negative definite if and only if it is < 0. Notice that another way of expressing this is to express it as a determinant of a 3×3 matrix:

$$-\left(a\beta^2 - 2b\alpha\beta + c\alpha^2\right) = \det \begin{pmatrix} 0 & \alpha & \beta \\ \alpha & a & b \\ \beta & b & c \end{pmatrix}$$

This matrix is constructed by "bordering" the 2×2 matrix A from the quadratic form of Q to the top and left by the coefficients of the linear constraint. So Q will be positive (negative) definite on the constraint space if the determinant of the matrix is negative (positive).

We can generalize this to $n$dimensions as follows. Consider a function Q of the form:

$$Q(x_1, ..., x_n) = \begin{pmatrix} x_1 & \cdots & x_n \end{pmatrix} \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = x'Ax$$

On the linear constraint set of m equations:

$$\begin{pmatrix} B_{11} & \cdots & B_{1n} \\ \vdots & \ddots & \vdots \\ B_{m1} & \cdots & B_{mn} \end{pmatrix} \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$Bx = 0$$

We can construct the $(m+n)\times(m+n)$ bordered matrix H as follows:

$$H = \left(\begin{array}{ccc|ccc} 0 & \cdots & 0 & B_{11} & \cdots & B_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & B_{m1} & \cdots & B_{mn} \\ \hline B_{11} & \cdots & B_{m1} & a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ B_{1n} & \cdots & B_{mn} & a_{1n} & \cdots & a_{nn} \end{array}\right) = \left(\begin{array}{c|c} 0 & B \\ \hline B' & A \end{array}\right)$$

Now we are ready for a theorem on how to determine the definiteness of a constrained quadratic form

Theorem 2. *To determine the definiteness of a n-variable quadratic form $Q(x) = x'Ax$ when restricted to the m-equation linear constraint set $Bx = 0$, we construct the $(m+n)\times(m+n)$ bordered matrix $H = \left(\begin{array}{c|c} 0 & B \\ \hline B' & A \end{array}\right)$ and check the following:*

(a) *If $\det(H)$ and the last $n-m$ leading principal minors all have the same sign as $(-1)^m$, then Q is positive definite on the constraint set*

(b) *If $\det(H)$ has the same sign as $(-1)^n$ and if the last $n-m$ leading principal minors alternate in sign, then Q is negative definite on the constraint set*

(c) *If both conditions (a) and (b) are violated by non-zero leading principal minors, then Q is indefinite on the constraint set*

Let's test our understanding with an example

Example. Check the definiteness of:

$$Q(x_1,x_2,x_3,x_4) = x_1^2 - x_2^2 + x_3^2 + x_4^2 + 4x_2x_3 - 2x_1x_4$$

On the constraint set:

$$\begin{aligned} x_2 + x_3 + x_4 &= 0 \\ x_1 - 9x_2 + x_4 &= 0 \end{aligned}$$

Let's write this is matrix notation: (note that $n = 4$ and $m = 2$)

$$\begin{aligned} Q(x) &= \sum_{i=1}^{4}\sum_{j=1}^{4} a_{ij}x_ix_j \\ &= \underbrace{a_{11}}_{=1} x_1^2 + \underbrace{(a_{14}+a_{41})}_{=-2} x_1x_4 + \underbrace{a_{22}}_{=1} x_2^2 + \underbrace{(a_{23}+a_{32})}_{=4} x_2x_3 + \underbrace{a_{33}}_{=1} x_3^2 + \underbrace{a_{44}}_{=1} x_4^2 \\ &= \begin{pmatrix} x_1 & x_2 & x_3 & x_4 \end{pmatrix} \begin{pmatrix} 1 & 0 & 0 & -1 \\ 0 & -1 & 2 & 0 \\ 0 & 2 & 1 & 0 \\ -1 & 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} \end{aligned}$$

$$Bx = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 1 & -9 & 0 & 1 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix}$$

Now we can construct the 6×6 bordered matrix H:

$$H = \left(\begin{array}{cc|cccc} 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & -9 & 0 & 1 \\ \hline 0 & 1 & 1 & 0 & 0 & -1 \\ 1 & -9 & 0 & -1 & 2 & 0 \\ 1 & 0 & 0 & 2 & 1 & 0 \\ 1 & 1 & -1 & 0 & 0 & 1 \end{array}\right)$$

So, we need to check the determinant of H and the last $n - m = 2$ leading principal minors (order 5 and 6). But notice that order 6 is just the determinant of H itself.

$$D_6 = \det(H) = 24$$

$$D_5 = \det \left(\begin{array}{cc|ccc} 0 & 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & -9 & 0 \\ \hline 0 & 1 & 1 & 0 & 0 \\ 1 & -9 & 0 & -1 & 2 \\ 1 & 0 & 0 & 2 & 1 \end{array}\right) = 77$$

Since $m = 2$, then $(-1)^m = 1$ is positive and so are D_5 and D_6. Therefore, Q is positive definite on the constraint set.

1.5 Convexity and Concavity

Recall that in Recitation 2 we saw that we could define convex and concave functions using the definiteness of their Hessian matrix.

Theorem 3. *Let $f : \mathbb{R}^n \Rightarrow \mathbb{R}$ be a C^2 function. Let $H(x)$ be the Hessian of $f(x)$. Then:*

- *f is concave $\Leftrightarrow$ $H(x)$ is negative semidefinite for all x*
- *f is convex $\Leftrightarrow$ $H(x)$ is positive semidefinite for all x*
- *$H(x)$ negative definite for all x $\Rightarrow$ f is strictly concave*
- *$H(x)$ positive definite for all x $\Rightarrow$ f is strictly convex*

Example. Determine whether $f(x_1, x_2, x_3) = -x_1^2 + 6x_1x_2 - 9x_2^2 - 2x_3^2$ is concave.

The gradient and Hessian are:

$$\nabla f(x) = \begin{pmatrix} -2x_1 + 6x_2 \\ -18x_2 + 6x_1 \\ -4x_3 \end{pmatrix}$$

$$H(x) = \begin{pmatrix} -2 & 6 & 0 \\ 6 & -18 & 0 \\ 0 & 0 & -4 \end{pmatrix}$$

We need to check if this is negative semidefinite. The leading principal minors are:

- Order 1: $\begin{vmatrix} -2 \end{vmatrix} = -2 < 0$ (delete {2,3})
- Order 2: $\begin{vmatrix} -2 & 6 \\ 6 & -18 \end{vmatrix} = 36 - 36 = 0$ (delete {3})
- Order 3: $\det(H) = -2 \cdot (72 - 0) - 6 \cdot (-24 - 0) + 0 \cdot (0 - 0) = -144 + 144 = 0$ (delete {})

The leading principal minor test is inconclusive (i.e. neither positive nor negative definite), so we need to check the remaining principal minors:

Order 1:

- $\begin{vmatrix} -18 \end{vmatrix} = -18 \leq 0$ (delete {1,3})
- $\begin{vmatrix} -4 \end{vmatrix} = -4 \leq 0$ (delete {1,2})

Order 2:

- $\begin{vmatrix} -18 & 0 \\ 0 & -4 \end{vmatrix} = 72 - 0 = 72 \geq 0$ (delete {1})
- $\begin{vmatrix} -2 & 0 \\ 0 & -4 \end{vmatrix} = 8 - 0 = 8 \geq 0$ (delete {2})

Clearly, this cannot be positive semidefinite (principal minors of order 1 are all negative). However, the conditions for negative semi-definite are satisfied (order 1 is ≤ 0, order 2 is ≥ 0, and order 3 is ≤ 0). Therefore, this function is concave.

In Recitation 2, I mentioned we could characterize quasi-convex/quasi-concave functions using a bordered Hessian. Now let's actually see this.

Theorem 4. *Let $f : S \to R$ be a C^2 function defined in an open, convex set $S \subseteq \mathbb{R}^n$. Let $f_i(x) = \frac{\partial f(x)}{\partial x_i}$ and $f_{ij}(x) = \frac{\partial^2 f(x)}{\partial x_j \partial x_i}$. Define the r^{th} order bordered Hessian as the $(r+1) \times (r+1)$ matrix:*

$$H_r(x) = \begin{pmatrix} 0 & f_1(x) & ... & f_r(x) \\ f_1(x) & f_{11}(x) & ... & f_{1r}(x) \\ \vdots & \vdots & \ddots & \vdots \\ f_r(x) & f_{r1}(x) & ... & f_{rr}(x) \end{pmatrix}$$

Moreover, define the determinant of this matrix as $B_r(x) = det(H_r(x))$ (i.e. the leading principal minor of order r, D_r, of the matrix H_n).

1. *If f is quasi-concave, then $(-1)^r B_r(x) \geq 0$, $\forall x \in S$ and all $r \in \{1, .., n\}$. Equivalently, $B_1(x) \leq 0, B_2(x) \geq 0, ...$ and $B_n(x) \leq 0$ if n is odd and $B_n(x) \geq 0$ if n is even,*
2. *If f is quasi-convex, then $B_r(x) \leq 0$, $\forall x \in S$ and all $r \in \{1, .., n\}$.*
3. *If $(-1)^r B_r(x) > 0$, $\forall x \in S$ and all $r \in \{1, .., n\}$, then f is quasi-concave. Equivalently, $B_1(x) < 0, B_2(x) > 0, ...$ and $B_n(x) < 0$ if n is odd and $B_n(x) > 0$ if n is even.*
4. *If $B_r(x) < 0$, $\forall x \in S$ and all $r \in \{1, .., n\}$, then f is quasi-convex.*

Why should we border the function's Hessian by the first derivatives? This suggests a constraint of $\sum_i f_i(x) x_i = \nabla f(x) \cdot x = 0$. The idea here is as follows. For a concave function, we need the Hessian to be negative semi-definite for all x. For a quasi-concave function, we only need the Hessian to be negative semi-definite for $\{x \in S : \nabla f(x) \cdot x = 0\}$. Geometrically, this says that we need the Hessian to be negative semi-definite on the tangent to the level curve of f through x. In other words, since quasi-concave functions have convex upper contour sets, the function is weakly decreasing along a vector tangent to a level curve.

Notice that the first two are necessary conditions and the last two are sufficient conditions. For example, if we find that a function has $B_r(x) \leq 0, \forall r$ and $B_r(x) = 0$ for some r and some x, then we cannot conclude the function is quasi-convex. It doesn't rule out the possibility, but we would need to check it in a different way.

Example. Consider the Cobb-Douglas utility function $u(x_1, x_2) = x_1^\alpha x_2^\beta$, with $\alpha, \beta > 0$. Show that this is quasi-concave

The bordered Hessian H is:

$$H = \begin{pmatrix} 0 & \alpha x_1^{\alpha-1} x_2^\beta & \beta x_1^\alpha x_2^{\beta-1} \\ \alpha x_1^{\alpha-1} x_2^\beta & \alpha(\alpha-1) x_1^{\alpha-2} x_2^\beta & \alpha\beta x_1^{\alpha-1} x_2^{\beta-1} \\ \beta x_1^\alpha x_2^{\beta-1} & \alpha\beta x_1^{\alpha-1} x_2^{\beta-1} & \beta(\beta-1) x_1^\alpha x_2^{\beta-2} \end{pmatrix}$$

Now let's calculate the leading principal minors:

$$\begin{aligned} B_1(x) =& -\left(\alpha x_1^{\alpha-1} x_2^\beta\right)^2 < 0 \\ B_2(x) =& 0 \left(\alpha(\alpha-1) x_1^{\alpha-2} x_2^\beta \cdot \beta(\beta-1) x_1^\alpha x_2^{\beta-2} - \alpha\beta x_1^{\alpha-1} x_2^{\beta-1} \cdot \alpha\beta x_1^{\alpha-1} x_2^{\beta-1}\right) \\ & - \alpha x_1^{\alpha-1} x_2^\beta \left(\alpha x_1^{\alpha-1} x_2^\beta \cdot \beta(\beta-1) x_1^\alpha x_2^{\beta-2} - \alpha\beta x_1^{\alpha-1} x_2^{\beta-1} \cdot \beta x_1^\alpha x_2^{\beta-1}\right) \end{aligned}$$

$$
\begin{aligned}
&+ \beta x_1^{\alpha} x_2^{\beta-1} \left(\alpha x_1^{\alpha-1} x_2^{\beta} \cdot \alpha\beta x_1^{\alpha-1} x_2^{\beta-1} - \alpha(\alpha-1) x_1^{\alpha-2} x_2^{\beta} \cdot \beta x_1^{\alpha} x_2^{\beta-1} \right) \\
=&0 + \underbrace{\left[-\alpha^2\beta(\beta-1) + \alpha^2\beta^2\right]}_{>0} x_1^{3\alpha-2} x_2^{3\beta-2} + \underbrace{\left[\alpha^2\beta^2 - \alpha(\alpha-1)\beta^2\right]}_{>0} x_1^{3\alpha-2} x_2^{3\beta-2} > 0
\end{aligned}
$$

We can see that condition 3 is satisfied, and therefore, this function is quasi-concave.

2 Optimization

2.1 Important Results

The concepts we have covered will be very useful for optimization. For a C^1 function $f : S \to \mathbb{R}$, let's define a few more terms:

Max/Min $x^* \in S$ is a (global) max of f if $f(x^*) \geq f(x), \forall x \in S$. Similarly, $x^* \in S$ is a (global) min of f if $f(x^*) \leq f(x), \forall x \in S$.

Strict Max $x^* \in S$ is a strict max of f if $f(x^*) > f(x), \forall x \in S \backslash \{x^*\}$.

Local Max $x^* \in S$ is a local max of f if there is an open ball $B(x^*, r)$ such that $f(x^*) \geq f(x), \forall x \in B(x^*, r) \cap S$.

Interior Point $x^* \in S$ is an interior point of f if there is an ε-open ball around x^* in the domain of f: $B(x^*, \varepsilon) \subset S$

Critical Point $x^* \in S$ is a critical point of f if $\frac{\partial f(x^*)}{\partial x_i} = 0, \forall i$, i.e. $\nabla f(x^*) = 0$

Now we are ready to go through the results - be careful to note which is a necessary or sufficient condition. For all of these, assume $f : S \to \mathbb{R}$ is a C^2 function on an open set $S \subset \mathbb{R}^n$. For a multi-variable function, denote the gradient as $\nabla f(x)$ and the Hessian matrix as $D^2 f(x)$.

For the the first set of results, let's look at conditions for a local max/min of an unconstrained function.

Theorem 5. ***Unconstrained One Variable Function: Necessary, Local***

- If x^* is a local max and an interior point of $f(x)$, then $f(x^*) = 0$ and $f''(x^*) \leq 0$
- If x^* is a local min and an interior point of $f(x)$, then $f(x^*) = 0$ and $f''(x^*) \geq 0$

Theorem 6. ***Unconstrained One Variable Function: Sufficient, Local***

- If $f'(x^*) = 0$ and $f''(x^*) < 0$, x^* is a strict local max of f
- If $f'(x^*) = 0$ and $f''(x^*) > 0$, x^* is a strict local min of f

Theorem 7. ***Unconstrained Multi-Variable Function: Necessary, Local***

- If x^* is a local max and an interior point of $f(x)$, then $\nabla f(x^*) = 0$ and $D^2 f(x^*)$ is negative semidefinite
- If x^* is a local min and an interior point of $f(x)$, then $\nabla f(x^*) = 0$ and $D^2 f(x^*)$ is positive semidefinite

Theorem 8. ***Unconstrained Multi-Variable Function: Sufficient, Local***

- If $\nabla f(x^*) = 0$ and $D^2 f(x^*)$ is negative definite, then x^* is a strict local max of f
- If $\nabla f(x^*) = 0$ and $D^2 f(x^*)$ is positive definite, then x^* is a strict local min of f
- If $\nabla f(x^*) = 0$ and $D^2 f(x^*)$ is indefinite, then x^* is neither a local max or min of f

For the the second set of results, let's look at conditions for a local max/min of constrained function. In particular, define the constraint set to be a set of m equations: $C_g = \{x \in S : g_1(x) = 0, ..., g_m(x) = 0\}$. Therefore, the Lagrangian is:

$$\mathcal{L}(x, \lambda) = f(x) + \sum_{j=1}^{m} \lambda_j g_j(x)$$

And define the bordered Hessian as:

$$H(x, \lambda) = \begin{pmatrix} 0 & Dg(x) \\ Dg(x) & D_x^2 \mathcal{L}(x, \lambda) \end{pmatrix}$$

Where $Dg(x)$ is the Jacobian of the function $g(x) : \mathbb{R}^n \to \mathbb{R}^m$ defined as $g(x) = [g_1(x), ..., g_m(x)]^T$.

Theorem 9. ***Constrained Multi-Variable Function: Sufficient, Local***
For an $x^* \in C_g$ *and* $\lambda^* = (\lambda_1^*, ..., \lambda_m^*)$ *such that* $\frac{\partial \mathcal{L}}{\partial x_i} = 0, \forall i$ *at* $(x_1^*, ..., x_n^*, \lambda_1^*, ..., \lambda_m^*)$

- If $D_x^2 \mathcal{L}(x, \lambda)$ is negative definite on the constraint set, then x^* is a strict local max of f on the constraint set C_g. Equivalently, if $H(x, \lambda)$ satisfies the conditions of Theorem 2(b).
- If $D_x^2 \mathcal{L}(x, \lambda)$ is positive definite on the constraint set, then x^* is a strict local min of f on the constraint set C_g. Equivalently, if $H(x, \lambda)$ satisfies the conditions of Theorem 2(a).

Theorem 10. ***One Constraint, Two-Variable Function: Sufficient, Local***[1]
For an $x^* \in C_g$ *and* λ^* *such that* $\frac{\partial \mathcal{L}}{\partial x_1} = 0$ *and* $\frac{\partial \mathcal{L}}{\partial x_2} = 0$ *at* $(x_1^*, x_2^*, \lambda^*)$

- If $\det(H(x^*, \lambda^*)) > 0$, then x^* is a strict local max of f on the constraint set C_g
- If $\det(H(x^*, \lambda^*)) < 0$, then x^* is a strict local min of f on the constraint set C_g

Notice that all these conditions are about local maximizers and minimizers. If we want to make claims about a global max/min, it is really handy to know about the concavity of the function. For the next two results, assume S is a convex open subset of $\mathbb{R}^n$.

Theorem 11. ***Unconstrained Multi-Variable Function: Sufficient, Global***

- If f is a concave function on S (i.e. $D^2 f(x)$ is negative semidefinite $\forall x$) and $\nabla f(x^*) = 0$, then x^* is a global max of f. Moreover, if f is strictly concave, then x^* is a the unique global max.
- If f is a convex function on S (i.e. $D^2 f(x)$ is positive semidefinite $\forall x$) and $\nabla f(x^*) = 0$, then x^* is a global min of f. Moreover, if f is strictly convex, then x^* is a the unique global min.

Theorem 12. ***Unconstrained Multi-Variable Function: Sufficient, Global***

- If f is a quasi-concave function on S and $\nabla f(x^*) = 0$ and $D^2 f(x^*)$ is negative definite, then x^* is a global max of f. Moreover, if f is strictly quasi-concave, then x^* is a the unique global max.
- If f is a quasi-convex function on S and $\nabla f(x^*) = 0$ and $D^2 f(x^*)$ is positive definite, then x^* is a global min of f. Moreover, if f is strictly quasi-convex, then x^* is a the unique global min.

[1] Note, we are just taking the above with $n = 2$ (two variables) and $m = 1$ (one equation constraint). This means we only need to look at the last $n - m = 1$ leading principal minors. But this means just looking at the determinant of the bordered Hessian itself. Moreover, for positive definite, we need the determinate to have the same sign as $(-1)^m = -1$. For negative definite, we need the determinate to have the same sign as $(-1)^n = 1$.

2.2 Examples

Example. Determine whether the critical point of the following system is a local max or min:

$$\max_{x,y} y^2 - x^2 \text{ s.t. } \alpha x + y = 3$$

We don't actually need to find the critical point, since the bordered Hessian is constant:

$$D = \begin{vmatrix} 0 & \alpha & 1 \\ \alpha & -2 & 0 \\ 1 & 0 & 2 \end{vmatrix} = -2\alpha^2 + 2,$$

Note that we are in the two-variable, one constraint case ($n = 2$,$m = 1$). Thus if $|\alpha| > 1$, $D < 0$, it will be a local min. If $|\alpha| < 1$, $D > 0$, it will be a local max.

Example. Maximize $x^2y^2z^2$ subject to the constraint $x^2 + y^2 + z^2 = 3$

The Lagrangian is:

$$\mathcal{L}(x, y, z, \lambda) = x^2y^2z^2 + \lambda\left(3 - x^2 - y^2 - z^2\right)$$

The FOCs are:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial x} &= 2xy^2z^2 - 2\lambda x = 0 \\ \frac{\partial \mathcal{L}}{\partial y} &= 2x^2yz^2 - 2\lambda y = 0 \\ \frac{\partial \mathcal{L}}{\partial z} &= 2x^2y^2z - 2\lambda z = 0 \\ \frac{\partial \mathcal{L}}{\partial \lambda} &= 3 - x^2 - y^2 - z^2 = 0 \end{aligned}$$

Dividing the first FOC by the second:

$$\begin{aligned} \frac{2xy^2z^2}{2x^2yz^2} &= \frac{2\lambda x}{2\lambda y} \\ \frac{y}{x} &= \frac{x}{y} \\ \implies y^2 &= x^2 \end{aligned}$$

The same could be done with the third, which tells us $x^2 = y^2 = z^2$. Therefore:

$$\begin{aligned} \frac{\partial \mathcal{L}}{\partial \lambda} &= 3 - x^2 - y^2 - z^2 \\ 0 &= 3 - 3x^2 \\ \implies 1 &= x^2 = y^2 = z^2 \end{aligned}$$

Moreover, plugging this into the first FOC (or any of the first three) gives us: $2x(y^2z^2 - \lambda) = 2x(1 - \lambda) = 0 \implies \lambda = 1$. Therefore, the possible solutions are $(x^*, y^*, z^*, \lambda^*) = (\pm 1, \pm 1, \pm 1, 1)$.

Next, let's check the SOC. Since $n = 3$ and $m = 1$, the bordered Hessian matrix is a 4×4 matrix:

$$H(x, y, z, \lambda) = \begin{pmatrix} 0 & -2x & -2y & -2z \\ -2x & 2y^2z^2 - 2\lambda & 4xyz^2 & 4xy^2z \\ -2y & 4xyz^2 & 2x^2z^2 - 2\lambda & 4x^2yz \\ -2z & 4xy^2z & 4x^2yz & 2x^2y^2 - 2\lambda \end{pmatrix}$$

Since $n = 3$ and $m = 1$, we are going to have to consider the last $n - m = 2$ leading principal minors (i.e. orders 3 and 4). Consider $(-1, -1, -1)$, then the bordered matrix becomes:

$$H(-1, -1, -1, 1) = \begin{pmatrix} 0 & 2 & 2 & 2 \\ 2 & 0 & 4 & 4 \\ 2 & 4 & 0 & 4 \\ 2 & 4 & 4 & 0 \end{pmatrix}$$

The leading principal minor of order 3 is $\begin{vmatrix} 0 & 2 & 2 \\ 2 & 0 & 4 \\ 2 & 4 & 0 \end{vmatrix} = 32$.

The leading principal minor of order 4 is $|H| = -192$.

As we can see, these alternate in sign with $|H|$ having the same sign as $(-1)^n = (-1)^3 = -1$. Therefore, the Hessian is negative definite on the constraint set, which means this point is indeed a strict local max on the constraint set.