7 Extremal Points and Higher Derivatives

2022-23

Recall that in Chapter 5 we looked at differentiable $f : \mathbb{R}^n \to \mathbb{R}$ and asked for the extremal points of f restricted to a surface $\{\mathbf{x} : \mathbf{g}(\mathbf{x}) = 0\}$ for some $\mathbf{g} : \mathbb{R}^n \to \mathbb{R}^m$. When the extremal points had been found, for example by Lagrange's method, we then usually used observation to see if we had a minimum or maximum or neither.

In this section, with $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$, we look for extremal values of f in U. Here U is not a surface within $\mathbb{R}^n$ but an open subset. As noted in Chapter 5 these extremal points are a subset of the critical points which are easy to find. Much more difficult, and the subject of this section, is to find a method which indicates whether a given critical point is a minimum or maximum or neither.

For functions of one variable there is a well-known test for the nature of a critical point which is given by the sign of the second derivative. Assuming $f'(a) = 0$ if $f''(a) > 0$ then f has a minimum at a , while if $f''(a) < 0$ then f has a maximum at a. This result can be generalised to functions of severable variables but to do this we have to consider higher derivatives.

7.1 The Hessian Matrix

The student should look back to the Chapter on Differential Forms to find the definitions of higher derivatives of functions of several variables and of $\mathcal{C}^q$-functions. The important result from that section was

Theorem 1 *If a function $f : U \to \mathbb{R}$ where U is open in $\mathbb{R}^n$ is of class $\mathcal{C}^2$ then*

$$\frac{\partial^2 f}{\partial x^j \partial x^i} = \frac{\partial^2 f}{\partial x^i \partial x^j}$$

for all i and j.

The important object when looking at extrema is

Definition 2 *Given a scalar-valued function $f : U \to \mathbb{R}$ of class $\mathcal{C}^2$, where U is open in $\mathbb{R}^n$, we define the **Hessian matrix** of f at $\mathbf{a} \in U$ to be the $n \times n$ matrix $Hf(\mathbf{a})$ with (i,j)-th entry $\partial^2 f(\mathbf{a}) / \partial x^i \partial x^j$.*

This matrix is symmetric by Theorem 1.

Example 3 *Let $f(\mathbf{x}) = x^2 y + y^2 z + z^2 - 2x$ for $\mathbf{x} \in \mathbb{R}^3$. The Hessian matrix is*

$$Hf(\mathbf{x}) = \begin{pmatrix} 2y & 2x & 0 \\ 2x & 2z & 2y \\ 0 & 2y & 2 \end{pmatrix}.$$

■

7.2 Taylor's Theorem

The Hessian matrix occurs in a natural way within the next important result. Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ be a scalar-valued function and $\mathbf{a} \in U$. Let $\mathbf{h} \in \mathbb{R}^n$ such that $\mathbf{a} + \mathbf{h} \in U$. Define the function $\phi : \mathbb{R} \to \mathbb{R}$ by $\phi(t) = f(\mathbf{a} + t\mathbf{h})$ for $0 \le t \le 1$. Then Taylor's Theorem for real-valued, functions of one variable with continuous derivative implies that, with Lagrange's form of the error,

$$\phi(t) = \phi(0) + \phi'(0)\, t + \phi''(c)\, \frac{t^2}{2}, \tag{1}$$

for some $0 < c < t$. The first derivative in (1) is a simple application of the Chain Rule:

$$\phi'(0) = \left. \frac{d}{dt} f(\mathbf{a} + t\mathbf{h}) \right|_{t=0} = \nabla f(\mathbf{a}) \bullet \mathbf{h}.$$

For the second derivative in (1)

$$\phi''(c) = \left. \frac{d}{dt} \phi'(t) \right|_{t=c} = \left. \frac{d}{dt} \nabla f(\mathbf{a} + t\mathbf{h}) \bullet \mathbf{h} \right|_{t=c}$$

Here

$$\frac{d}{dt} \nabla f(\mathbf{a} + t\mathbf{h}) \bullet \mathbf{h} = \sum_{i=1}^{n} h^i \frac{d}{dt} \frac{\partial f}{\partial x^i} (\mathbf{a} + t\mathbf{h}).$$

The Chain Rule again gives

$$\frac{d}{dt} \frac{\partial f}{\partial x^i} (\mathbf{a} + t\mathbf{h}) = \sum_{j=1}^{n} \frac{\partial^2 f}{\partial x^j \partial x^i} (\mathbf{a} + t\mathbf{h})\, h^j.$$

Substituting back,

$$\phi''(c) = \sum_{i=1}^{n}\sum_{j=1}^{n} h^i h^j \frac{\partial^2 f}{\partial x^j \partial x^i}(\mathbf{a} + c\mathbf{h}) = \mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h},$$

where $\mathbf{c} = \mathbf{a} + c\mathbf{h}$. Return to (1) with $t = 1$ to find

Theorem 4 ***Taylor's Theorem*** *Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ be a function of class $\mathcal{C}^2$ and $\mathbf{a} \in U$. Let $\mathbf{h} \in \mathbb{R}^n$such that $\mathbf{a} + \mathbf{h} \in U$. Then*

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + \nabla f(\mathbf{a}) \bullet \mathbf{h} + \frac{1}{2}\mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h}, \tag{2}$$

where $\mathbf{c} = \mathbf{a} + c\mathbf{h}$ for some $c : 0 \le c \le 1$.

7.3 Graphs of dimension n within $\mathbb{R}^{n+1}$

. Given a scalar-valued $\mathcal{C}^2$-function $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ then the graph

$$S = \left\{ \begin{pmatrix} \mathbf{x} \\ f(\mathbf{x}) \end{pmatrix} : \mathbf{x} \in U \right\} \subseteq \mathbb{R}^{n+1}$$

is a surface as long as $Jf(\mathbf{x})$ is well-defined for all $\mathbf{x} \in U$ as we now assume. Let $\mathbf{p} \in S$ so

$$\mathbf{p} = \begin{pmatrix} \mathbf{q} \\ f(\mathbf{q}) \end{pmatrix}$$

for some $\mathbf{q} \in U$. We can express S around $\mathbf{p}$ as

$$\left\{ \begin{pmatrix} \mathbf{q} + \mathbf{h} \\ f(\mathbf{q} + \mathbf{h}) \end{pmatrix} : \mathbf{h} \in U_{\mathbf{q}} \right\} = \left\{ \begin{pmatrix} \mathbf{q} + \mathbf{h} \\ f(\mathbf{q}) + Jf(\mathbf{q})\,\mathbf{h} + \frac{1}{2}\mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h} \end{pmatrix} : \mathbf{h} \in U_{\mathbf{q}} \right\}. \tag{3}$$

where $U_{\mathbf{q}} = U - \mathbf{q}$. This can be compared with Tangent plane to S at $\mathbf{p}$,

$$\mathbf{p} + T_{\mathbf{p}}(S) = \left\{ \begin{pmatrix} \mathbf{q} + \mathbf{h} \\ f(\mathbf{q}) + Jf(\mathbf{q})\,\mathbf{h} \end{pmatrix} : \mathbf{h} \in \mathbb{R}^n \right\}.$$

The Tangent Space is of dimension n inside $\mathbb{R}^{n+1}$ and so has only one normal direction, $\mathbf{N} \in \mathbb{R}^{n+1}$, say. Being normal to $T_{\mathbf{p}}(S)$ means that $\mathbf{N} \bullet \mathbf{w} = 0$ for all $\mathbf{w} \in T_{\mathbf{p}}(S)$. That is

$$\mathbf{N} \bullet \begin{pmatrix} \mathbf{h} \\ Jf(\mathbf{q})\,\mathbf{h} \end{pmatrix} = 0 \tag{4}$$

for all $\mathbf{h} \in \mathbb{R}^n$.

If you are on the surface (3) at $\mathbf{p}$ and move away then

$$\mathbf{p} \to \mathbf{p} + \begin{pmatrix} \mathbf{h} \\ Jf(\mathbf{q})\,\mathbf{h} + \frac{1}{2}\mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h} \end{pmatrix}.$$

The component of this change in the normal direction is

$$\begin{aligned} \mathbf{N} \bullet \begin{pmatrix} \mathbf{h} \\ Jf(\mathbf{q})\,\mathbf{h} + \frac{1}{2}\mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h} \end{pmatrix} &= \mathbf{N} \bullet \begin{pmatrix} \mathbf{h} \\ Jf(\mathbf{q})\,\mathbf{h} \end{pmatrix} + \mathbf{N} \bullet \begin{pmatrix} \mathbf{0} \\ \frac{1}{2}\mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h} \end{pmatrix} \\ &= 0 + N\frac{1}{2}\mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h}, \end{aligned}$$

by (4). (Here N is the $n{+}1$-th coordinate of $\mathbf{N}$.)

If $\mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h} \geq 0$ for all $|\mathbf{h}| < \delta$ with $\delta > 0$ sufficiently small, then the component of the surface (3) in the normal direction has the same sign for all such $\mathbf{h}$. This means that locally to $\mathbf{p}$ the surface stays to one side of the Tangent plane. Similarly if $\mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h} \leq 0$ for all sufficiently small $\mathbf{h}$.

If, for all $\delta > 0$ there exists $|\mathbf{h}_1|, |\mathbf{h}_2| < \delta : \mathbf{h}_1^T Hf(\mathbf{c})\,\mathbf{h}_1 > 0$ and $\mathbf{h}_2^T Hf(\mathbf{c})\,\mathbf{h}_2 < 0$ for some $\mathbf{h}_2 \in \mathbb{R}^n$ then locally the surface has parts on both sides of the Tangent plane.

Though we appear to be getting interesting information from the Hessian matrix it is difficult to apply this in practice since $\mathbf{c}$ depends on $\mathbf{h}$. But, f is a function of class $\mathcal{C}^2$, and so $Hf(\mathbf{x})$ is a continuous function of $\mathbf{x}$. Thus, for $\mathbf{h}$ sufficiently small $\mathbf{c}$ will be close to $\mathbf{q}$ and $\mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h}$ close in value to $\mathbf{h}^T Hf(\mathbf{q})\,\mathbf{h}$. In particular if one is positive then so is the other. Hence, by looking at the sign of $\mathbf{h}^T Hf(\mathbf{q})\,\mathbf{h}$ for small $\mathbf{h}$ we can see whether a surface lies, locally to one side of the tangent plane or not. We give a name to the properties of Hessian matrix we are looking for.

Definition 5 *Assume A is a symmetric real matrix. We say that*

A is ***positive definite*** *iff $\mathbf{v}^T A\mathbf{v} > 0$ for all non-zero vectors $\mathbf{v}$,*

A is ***negative definite*** *iff $\mathbf{v}^T A\mathbf{v} < 0$ for all non-zero vectors $\mathbf{v}$,*

A is ***indefinite*** *if these exists $\mathbf{v} \neq \mathbf{0} : \mathbf{v}^T A\mathbf{v} > 0$ and there exists $\mathbf{v} \neq \mathbf{0} : \mathbf{v}^T A\mathbf{v} < 0$,*

otherwise A is ***nondefinite****.*

Note the nondefinite definition covers the cases such as $\mathbf{v}^T A\mathbf{v} \geq 0$ for all $\mathbf{v}$ with $\mathbf{v}^T A\mathbf{v} = 0$ for some *non-zero* vector $\mathbf{v}$.

If $Hf(\mathbf{q})$ is definite then locally S remains to one side of the Tangent plane while If $Hf(\mathbf{q})$ is indefinite then, arbitrarily close to $\mathbf{p}$, S has parts on both sides of the Tangent Plane .

7.4 Extremum points

Look back at the start of Chapter 5 to find the definitions for local $\backslash$ strict $\backslash$ maximum $\backslash$ minimum $\backslash$ extremum. The important result there was that if $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ had a local extremum at $\mathbf{a}$ then $\nabla f(\mathbf{a}) = 0$, i.e. $\mathbf{a}$ is a critical point.

Definition 6 *A critical point* $\mathbf{a} \in U$ *of* f *is* ***non-degenerate*** *when the Hessian matrix of* f *at* $\mathbf{a}$ *is non-singular (i.e. invertible).*

When a critical point is non-degenerate, the nature of the critical point is determined by the Hessian matrix. Return to Taylor's Theorem which, at a critical point $\mathbf{a}$, states that for all $\mathbf{h} : \mathbf{a} + \mathbf{h} \in U$,

$$f(\mathbf{a} + \mathbf{h}) = f(\mathbf{a}) + \frac{1}{2}\mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h},$$

where $\mathbf{c} = \mathbf{a} + c\mathbf{h}$ for some $c : 0 \leq c \leq 1$.

From this we see that $f(\mathbf{a})$ is a *local* strict minimum iff $f(\mathbf{a} + \mathbf{h}) > f(\mathbf{a})$ iff $\mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h} > \mathbf{0}$ for all $\mathbf{h} \neq \mathbf{0}$ sufficiently small. By the observation above this is equivalent to $\mathbf{h}^T Hf(\mathbf{a})\,\mathbf{h} > \mathbf{0}$ for all $\mathbf{h} \neq \mathbf{0}$ sufficiently small

Similarly, $f(\mathbf{a})$ is a *local* strict maximum iff $\mathbf{h}^T Hf(\mathbf{a})\,\mathbf{h} < \mathbf{0}$ for all $\mathbf{h} \neq \mathbf{0}$ sufficiently small.

Combined with Definition 5 and we have

If $Hf(\mathbf{a})$ is positive definite then f has a *local strict minimum* at the critical point $\mathbf{a}$. If $Hf(\mathbf{a})$ is negative definite then f has a *local strict maximum* at the critical point $\mathbf{a}$. If $Hf(\mathbf{a})$ is indefinite then we say that f has a *saddle* at the critical point $\mathbf{a}$.

Why do we say we have a saddle in the last case?

Perhaps look at the graph of f in $\mathbb{R}^{n+1}$. The Tangent plane at a critical point $\mathbf{a}$ is

$$\left\{ \begin{pmatrix} \mathbf{a}+\mathbf{h} \\ f(\mathbf{a}) \end{pmatrix} : \mathbf{h} \in \mathbb{R}^n \right\}.$$

So as we move away from the point $\mathbf{p} = (\mathbf{a}, f(\mathbf{a}))^T$ the height, i.e. the n+1-th coordinate, does not change. But around $\mathbf{p}$ the surface can be written as

$$\left\{ \begin{pmatrix} \mathbf{a}+\mathbf{h} \\ f(\mathbf{a}) + \frac{1}{2}\mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h} \end{pmatrix} : \mathbf{h} \in \mathbb{R}^n \right\}.$$

If we find an $\mathbf{h}_1 : \mathbf{h}_1^T Hf(\mathbf{c})\,\mathbf{h}_1 > \mathbf{0}$ (equivalent to $\mathbf{h}_1^T Hf(\mathbf{a})\,\mathbf{h}_1 > \mathbf{0}$) then the last coordinate increases as we move away in that direction. If we find an $\mathbf{h}_2 : \mathbf{h}_2^T Hf(\mathbf{a})\,\mathbf{h}_2 < \mathbf{0}$ then the last coordinate decreases as we move away in that direction. This is a property of a saddle.

Return now to

Example 3 continued The critical values of $f(\mathbf{x}) = x^2y + y^2z + z^2 - 2x$ are solutions of

$$\mathbf{0} = \nabla f(\mathbf{x}) = \begin{pmatrix} 2xy - 2 \\ x^2 + 2yz \\ y^2 + 2z \end{pmatrix}.$$

That is, $xy = 1$, $x^2 + 2yz = 0$ and $y^2 + 2z = 0$. Multiply the second equation by x and use the first to get $x^3 + 2z = 0$. Subtract the third equation to find $x^3 = y^2$. Raise the first equation by the third power and use the result just found to get $y^5 = 1$. Thus $y = \pm 1$. In $x^3 = y^2$ and we get $x = 1$. But then $xy = 1$ implies $y = 1$. Finally $y^2 + 2z = 0$ gives $z = -1/2$. So there is only one critical point $\mathbf{a} = (1, 1, -1/2)^T$.

The Hessian matrix at this critical point is

$$Hf(\mathbf{a}) = \begin{pmatrix} 2 & 2 & 0 \\ 2 & -1 & 2 \\ 0 & 2 & 2 \end{pmatrix}.$$

The determinant is -20, i.e. non-zero, and so $Hf(\mathbf{a})$ is non-singular and hence $\mathbf{a}$ is non-degenerate.

Next observe that

$$\mathbf{e}_1^T Hf(\mathbf{a})\,\mathbf{e}_1 = \begin{pmatrix} 1 & 0 & 0 \end{pmatrix} \begin{pmatrix} 2 & 2 & 0 \\ 2 & -1 & 2 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} = 2 > 0$$

while

$$\mathbf{e}_2^T Hf(\mathbf{a})\,\mathbf{e}_2 = \begin{pmatrix} 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 2 & 2 & 0 \\ 2 & -1 & 2 \\ 0 & 2 & 2 \end{pmatrix} \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix} = -1 < 0.$$

Thus $Hf(\mathbf{a})$ is indefinite and $\mathbf{a}$ is a saddle point. ■

Unfortunately this method requires finding two vectors with different signs for $\mathbf{v}^T Hf(\mathbf{a})\,\mathbf{v}$. Is there a method which does not involve a search?

Tests for Definiteness or otherwise.

Question 1, Given a symmetric matrix A under what conditions is $\mathbf{v}^T A\mathbf{v} > 0$ for all non-zero vectors $\mathbf{v}$? When is $\mathbf{v}^T A\mathbf{v} < 0$ for all non-zero vectors $\mathbf{v}$?

Question 2 Given a symmetric matrix A, if there exist vectors $\mathbf{v}$ such that $\mathbf{v}^T A\mathbf{v} > 0$ and $\mathbf{w}$ such that $\mathbf{w}^T A\mathbf{w} < 0$ how can they be found?

The answer to Question 1 is easily given if A is a 2×2 matrix.

Theorem 7 *Suppose*

$$M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

If $\det M > 0$ *then* M *is positive definite when* $a > 0$ *and negative definite when* $a < 0$. *If* $\det M < 0$, *then* M *is indefinite. If* $\det M = 0$ *then* M *is nondefinite.*

Proof Exercise, but all results when $a \neq 0$ follow from

$$(x, y) \begin{pmatrix} a & b \\ b & c \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = a \left(x + \frac{by}{a} \right)^2 + \frac{\det A}{a} y^2.$$

See Appendix ■

Notation To ease congestion in the writing we already have

$$d_j f(\mathbf{x}) = \frac{\partial f}{\partial x^j}(\mathbf{x}).$$

Extend this to higher derivatives by

$$d_{i,j} f(\mathbf{x}) = \frac{\partial^2 f}{\partial x^j \partial x^i}(\mathbf{x}).$$

I've no idea why the order of the subscripts in $d_{i,j}$ is the reverse of the superscripts in $\partial^2 f / \partial x^j \partial x^i$ but just remember that it is so.

Example 8 *Find the critical points of*

$$f(\mathbf{x}) = \frac{1}{3}x^3 - 3x^2 + \frac{1}{4}y^2 + xy + 13x - y + 2,$$

for $\mathbf{x} \in \mathbb{R}^2$, *and find if they are local minima, maxima or saddle points.*

Solution The gradient vector is

$$\nabla f(\mathbf{x}) = \left(x^2 - 6x + y + 13,\ \frac{y}{2} + x - 1\right)^T.$$

Then $\nabla f(\mathbf{x}) = \mathbf{0}$ becomes

$$x^2 - 6x + y + 13 = 0 \quad \text{and} \quad 2x + y - 2 = 0.$$

Substitute $y = -2x + 2$ into the first equation to get $x^2 - 8x + 15 = 0$ which factorises as $(x - 3)(x - 5) = 0$. This gives two critical points

$$\mathbf{a}_1 = (3, -4)^T \quad \text{and} \quad \mathbf{a}_2 = (5, -8)^T.$$

The Hessian matrix is

$$Hf(\mathbf{x}) = \begin{pmatrix} 2x - 6 & 1 \\ 1 & 1/2 \end{pmatrix}.$$

Then

$$Hf(\mathbf{a}_1) = \begin{pmatrix} 0 & 1 \\ 1 & 1/2 \end{pmatrix} \quad \text{and} \quad Hf(\mathbf{a}_2) = \begin{pmatrix} 4 & 1 \\ 1 & 1/2 \end{pmatrix}.$$

By Theorem 7 $\det Hf(\mathbf{a}_1) = -1 < 0$ implies that the matrix $Hf(\mathbf{a}_1)$ is indefinite and we have a saddle at $\mathbf{a}_1$. Similarly $\det Hf(\mathbf{a}_2) = 1 > 0$ and $4 > 0$ together imply that the matrix $Hf(\mathbf{a}_2)$ is positive definite and we have a local minimum at $\mathbf{a}_2$. ∎

The situation is more complicated for larger n.

Linear Algebra. The fundamental result for symmetric matrices is the following.

Theorem 9 *If A is a real symmetric $n \times n$ matrix then $\mathbb{R}^n$ has an orthonormal basis of eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ satisfying $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$; the λ_i are the eigenvalues of A.*

This can be applied immediately to the Hessian matrix evaluated at a non-degenerate critical point, $Hf(\mathbf{a})$. For this is a real symmetric matrix and so has n eigenvalues. If $\mathbf{w}$ is an eigenvector with associated eigenvalue λ then

$$\mathbf{w}^T Hf(\mathbf{a})\,\mathbf{w} = \lambda \mathbf{w}^T \mathbf{w} = \lambda\,|\mathbf{w}|^2\,.$$

If we can find eigenvalues $\lambda_1 > 0$ and $\lambda_2 < 0$ this means that as we approach $\mathbf{a}$ along the line of direction of the first eigenvector $\mathbf{w}_1$ we have $\mathbf{w}_1^T Hf(\mathbf{a})\,\mathbf{w}_1 > 0$ (a local minimum) while approaching along the second eigenvector $\mathbf{w}_2$ will give $\mathbf{w}_2^T Hf(\mathbf{a})\,\mathbf{w}_2 < 0$ (a local maximum). We would then have a saddle point. This answers a question above of how to choose the directions when examining a critical point.

Example 3 continued For $f(\mathbf{x}) = x^2y + y^2z + z^2 - 2x$ the Hessian matrix the critical point $\mathbf{a} = (1, 1, -1/2)^T$ is

$$\begin{pmatrix} 2 & 2 & 0 \\ 2 & -1 & 2 \\ 0 & 2 & 2 \end{pmatrix}.$$

The eigenvalues of this matrix are $2, 1/2 + \sqrt{41}/2$ and $1/2 - \sqrt{41}/2$. Two are positive and one negative so $\mathbf{a}$ is a saddle point. ■

What can we say if the eigenvalues are all of the same sign?

From Linear Algebra we have

Proposition 10 *Let $A = (a_{ij})$ be a real symmetric $n \times n$ matrix.*

A is positive definite if, and only if, all eigenvalues are positive,

A is negative definite if, and only if, all eigenvalues are negative.

If some $\lambda_k > 0$ and some $\lambda_\ell < 0$ then A is indefinite.

Proof See Appendix. ■

Thus, looking back at the definition of positive and negative definition, we have

- all eigenvalues positive implies $\mathbf{v}^T Hf(\mathbf{a})\,\mathbf{v} > 0$ for all non-zero $\mathbf{v}$ which implies $\mathbf{a}$ is a local minimum,
- all eigenvalues negative implies $\mathbf{v}^T Hf(\mathbf{a})\,\mathbf{v} < 0$ for all non-zero $\mathbf{v}$ which implies $\mathbf{a}$ is a local maximum.

Example 8 revisited. For

$$f(\mathbf{x}) = \frac{1}{3}x^3 - 3x^2 + \frac{1}{4}y^2 + xy + 13x - y + 2,$$

the critical points are $\mathbf{a}_1 = (3, -4)^T$ and $\mathbf{a}_2 = (5, -8)^T$. The Hessian matrices at these points are

$$Hf(\mathbf{a}_1) = \begin{pmatrix} 0 & 1 \\ 1 & 1/2 \end{pmatrix} \quad \text{and} \quad Hf(\mathbf{a}_2) = \begin{pmatrix} 4 & 1 \\ 1 & 1/2 \end{pmatrix}.$$

The eigenvalues of $Hf(\mathbf{a}_1)$ are $1/4 + \sqrt{17}/4 > 0$ and $1/4 - \sqrt{17}/4 < 0$ so $\mathbf{a}$ is a saddle point.

The eigenvalues of $Hf(\mathbf{a}_2)$ are $9/4 + \sqrt{65}/4 > 0$ and $9/4 - \sqrt{65}/4 > 0$ so $\mathbf{a}$ is a local minimum. ■

This discussion about eigenvectors and eigenvalues is, in fact, a distraction since they are difficult to calculate and yet we don't need to. We only need to know the signs of the eigenvalues. Is there a simpler method to do this? Perhaps, but I leave that for the Appendix.

Appendix to Section 7

Quadratic Approximations

Taylors Theorem Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ be a function of class $\mathcal{C}^2$ and $\mathbf{a} \in U$. Let $\mathbf{h} \in \mathbb{R}^n$ such that $\mathbf{a} + \mathbf{h} \in U$. Then

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + Jf(\mathbf{a})\,\mathbf{h} + \frac{1}{2}\mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h},$$

where $\mathbf{c} = \mathbf{a} + c\mathbf{h}$ for some $c : 0 \le c \le 1$.

Thus the Quadratic approximation to f at $\mathbf{a}$ might be (replacing $\mathbf{c}$ by $\mathbf{a}$)

$$f(\mathbf{a}+\mathbf{h}) = f(\mathbf{a}) + Jf(\mathbf{a})\,\mathbf{h} + \frac{1}{2}\mathbf{h}^T Hf(\mathbf{a})\,\mathbf{h}.$$

To see how good this approximation is we first state

Lemma 11 *If* $M = \left(a^i_j\right) \in M_{n,n}(\mathbb{R})$ *then*

$$\left|\mathbf{t}^T M \mathbf{t}\right| < C\left|\mathbf{t}\right|^2$$

for all $\mathbf{t} \in \mathbb{R}^n$ *with* $C^2 = \sum_{i,j}\left(a^i_j\right)^2$.

Proof

$$\begin{aligned}
\mathbf{t}^T M \mathbf{t} &= \mathbf{t}^T \begin{pmatrix} \longleftarrow & \mathbf{r}^1 & \longrightarrow \\ \longleftarrow & \mathbf{r}^2 & \longrightarrow \\ & \vdots & \\ \longleftarrow & \mathbf{r}^n & \longrightarrow \end{pmatrix} \mathbf{t} = \left(t^1, t^2, ..., t^n\right) \begin{pmatrix} \mathbf{r}^1 \bullet \mathbf{t} \\ \mathbf{r}^2 \bullet \mathbf{t} \\ \\ \mathbf{r}^n \bullet \mathbf{t} \end{pmatrix} \\
&= \sum_{i=1}^n t^i \left(\mathbf{r}^i \bullet \mathbf{t}\right).
\end{aligned}$$

Then, by Cauchy-Schwarz,

$$\begin{aligned}
\left|\mathbf{t}^T M \mathbf{t}\right|^2 &= \left(\sum_{i=1}^n t^i \left(\mathbf{r}^i \bullet \mathbf{t}\right)\right)^2 \le \sum_{i=1}^n \left(t^i\right)^2 \sum_{i=1}^n \left(\mathbf{r}^i \bullet \mathbf{t}\right)^2 \\
&\le \left|\mathbf{t}\right|^2 \sum_{i=1}^n \left|\mathbf{r}^i\right|^2 \left|\mathbf{t}\right|^2,
\end{aligned}$$

by a second application of Cauchy-Schwarz. So the result follows with

$$C^2 = \sum_{i=1}^{n} \left|\mathbf{r}^i\right|^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \left(a^i_j\right)^2 .$$

$\blacksquare$

Proposition 12 *Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ be a function of class $\mathcal{C}^2$ and $\mathbf{a} \in U$. Then, for all $\varepsilon > 0$ there exists $\delta > 0$ such that $|\mathbf{c} - \mathbf{a}| < \delta$ implies*

$$\left|\mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h} - \mathbf{h}^T Hf(\mathbf{a})\,\mathbf{h}\right| < \varepsilon\,|\mathbf{h}|^2 .$$

Proof We apply Lemma 11 with

$$M = Hf(\mathbf{c}) - Hf(\mathbf{a})\,,$$

in which case the entries in the matrix M are

$$a^i_j = \frac{\partial^2}{\partial x^i \partial x^j} f(\mathbf{c}) - \frac{\partial^2}{\partial x^i \partial x^j} f(\mathbf{a})\,.$$

Let $\varepsilon > 0$ be given. Then, since f is a $\mathcal{C}^2$-class function, there exists $\delta_{i,j} > 0$ such that if $|\mathbf{c} - \mathbf{a}| < \delta_{i,j}$ then

$$\left|\frac{\partial^2}{\partial x^i \partial x^j} f(\mathbf{c}) - \frac{\partial^2}{\partial x^i \partial x^j} f(\mathbf{a})\right| < \frac{\varepsilon}{n}.$$

Let $\delta = \min_{i,j} \delta_{i,j}$ and assume c satisfies $|\mathbf{c} - \mathbf{a}| < \delta$ then the C of Lemma 11 satisfies

$$C^2 = \sum_{i=1}^{n}\sum_{j=1}^{n} \left(a^i_j\right)^2 \leq \sum_{i=1}^{n}\sum_{j=1}^{n} \left(\frac{\varepsilon}{n}\right)^2 = \varepsilon^2.$$

Hence our result follows. $\blacksquare$

Corollary 13 *Let $f : U \subseteq \mathbb{R}^n \to \mathbb{R}$ be a function of class $\mathcal{C}^2$ and $\mathbf{a} \in U$. Then*

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - Jf(\mathbf{a})\,\mathbf{h} - \mathbf{h}^T Hf(\mathbf{a})\,\mathbf{h}/2}{|\mathbf{h}|^2} = 0.$$

Proof Replacing $f(\mathbf{a}+\mathbf{h})$ by $f(\mathbf{a}) + Jf(\mathbf{a})\,\mathbf{h} + \mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h}/2$ we have

$$\frac{f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - Jf(\mathbf{a})\,\mathbf{h} - \mathbf{h}^T Hf(\mathbf{a})\,\mathbf{h}/2}{|\mathbf{h}|^2} = \frac{\mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h} - \mathbf{h}^T Hf(\mathbf{a})\,\mathbf{h}}{2\,|\mathbf{h}|^2}.$$

Let $\varepsilon > 0$ be given. By Proposition 12 there exists $\delta > 0$ such that if $|\mathbf{c}-\mathbf{a}| < \delta$ then

$$\frac{\left|\mathbf{h}^T Hf(\mathbf{c})\,\mathbf{h} - \mathbf{h}^T Hf(\mathbf{a})\,\mathbf{h}\right|}{2\,|\mathbf{h}|^2} < \frac{\varepsilon\,|\mathbf{h}|^2}{2\,|\mathbf{h}|^2} = \frac{\varepsilon}{2}. \tag{5}$$

Thus, if $|\mathbf{h}| < \delta$ then $\mathbf{c} = \mathbf{a} + c\mathbf{h}$ for some $0 < c < 1$ means $|\mathbf{c}-\mathbf{a}| < \delta$ and (5) gives

$$\left|\frac{f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - Jf(\mathbf{a})\,\mathbf{h} - \mathbf{h}^T Hf(\mathbf{a})\,\mathbf{h}/2}{|\mathbf{h}|^2}\right| < \varepsilon.$$

Hence we have verified the ε-δ definition of limit. ■

Lemma 14 *If M is a symmetric $n \times n$ matrix for which*

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{f(\mathbf{a}+\mathbf{h}) - f(\mathbf{a}) - Jf(\mathbf{a})\,\mathbf{h} - \mathbf{h}^T M\mathbf{h}/2}{|\mathbf{h}|^2} = 0, \tag{6}$$

then $M = Hf(\mathbf{a})$.

Proof If M_1 and M_2 are two symmetric matrices satisfying (6) then

$$\lim_{\mathbf{h}\to\mathbf{0}} \frac{\mathbf{h}^T M\mathbf{h}}{|\mathbf{h}|^2} = 0 \tag{7}$$

where $M = M_1 - M_2$.

First note that, for $1 \le i \le n$, we have $\mathbf{e}_i^T M\mathbf{e}_i = a_i^i$. So if we look at the directional limit, choosing $\mathbf{h} = h\mathbf{e}_i$ we find that

$$\frac{\mathbf{h}^T M\mathbf{h}}{|\mathbf{h}|^2} = \frac{h^2}{h^2} a_i^i = a_i^i.$$

Thus (7) implies the diagonal elements a_i^i are zero.

Next note that $\mathbf{e}_i^T M \mathbf{e}_j = a_j^i$ so

$$\begin{aligned}
(\mathbf{e}_i + \mathbf{e}_j)^T M (\mathbf{e}_i + \mathbf{e}_j) &= a_j^i + a_i^i + a_j^j + a_i^j \\
&= a_j^i + a_i^j \\
&\qquad \text{since the diagonal terms are zero} \\
&= 2a_j^i
\end{aligned}$$

since the matrix is symmetric. So setting $\mathbf{h} = h(\mathbf{e}_i + \mathbf{e}_j)$ we will deduce $a_j^i = 0$ for all i, j. Hence $M = 0$, i.e. $M_1 = M_2$.

■

In this sense $f(\mathbf{a}) + Jf(\mathbf{a})\,\mathbf{h} + \mathbf{h}^T Hf(\mathbf{a})\,\mathbf{h}/2$ is the *best* quadratic approximation to f at $\mathbf{a}$.

Matrices

A result stated without proof in the lectures.

Theorem 7 *Suppose*

$$M = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

If $\det M > 0$ *then* M *is positive definite when* $a > 0$ *and negative definite when* $a < 0$.

If $\det M < 0$, *then* M *is indefinite.*

If $\det M \neq 0$ *and* $a = 0$ *then* M *is indefinite if* $c \neq 0$ *and nondefinite if* $c = 0$.

If $\det M = 0$ *then* M *is nondefinite.*

Proof Writing $\mathbf{v} = (x, y)^T$, the binary form $\mathbf{v}^T A\mathbf{v}$ equals

$$
\begin{aligned}
(x,y)\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} &= ax^2 + 2bxy + cy^2 = a\left(x + \frac{by}{a}\right)^2 - \frac{b^2y^2}{a} + cy^2 \\
&\qquad \text{provided } a \neq 0, \\
&= a\left(x + \frac{by}{a}\right)^2 + \frac{ac - b^2}{a}y^2 \\
&= a\left(x + \frac{by}{a}\right)^2 + \frac{\det M}{a}y^2. \qquad (8)
\end{aligned}
$$

First assume that $\det M \neq 0$ along with $a \neq 0$.

i. If $\det M > 0$ and $a > 0$ then, from (8), $\mathbf{v}^T M\mathbf{v} \geq 0$ for all $\mathbf{v}$ and equals 0 only when $y = 0$ and $x + by/a = 0$ i.e. $x = 0$. That is $\mathbf{v}^T A\mathbf{v} = 0$ iff $\mathbf{v} = \mathbf{0}$. Hence M is positive definite.

ii. Similarly, if $\det M > 0$ and $a < 0$ then M is negative definite .

iii. If $\det M < 0$ then the coefficients of $(x + by/a)^2$ and y^2 are of different sign in which case M is indefinite.

Next assume that $\det M \neq 0$ and $a = 0$.

iv. If $\det M \neq 0$ and $a = 0$ and $c \neq 0$ (which combine to give $b \neq 0$) then

$$
(x,y)\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 2bxy + cy^2 = c\left(y + \frac{bx}{c}\right)^2 - \frac{b^2x^2}{c}.
$$

Whatever the sign of c the coefficients of $(y + bx/c)^2$ and b^2x^2 will be of different sign, in which case M is indefinite.

v. If $\det M \neq 0$ and $a = c = 0$ then

$$
(x,y)\begin{pmatrix} a & b \\ b & c \end{pmatrix}\begin{pmatrix} x \\ y \end{pmatrix} = 2bxy,
$$

which can be zero for non-zero vectors (i.e. $(0,1)\,M\,(0,1)^T = 0$) and so M is nondefinite.

vi. Finally, if $\det A = 0$ then $\mathbf{v}^T A\mathbf{v} = 0$ if $\mathbf{v} = (-b, a)$, say, in which case M is nondefinite. ■

Eigenvalues

The fundamental result for symmetric matrices is the following.

Theorem 15 *If A is a real symmetric $n \times n$ matrix then $\mathbb{R}^n$ has an orthonormal basis of eigenvectors $\{\mathbf{v}_1, \mathbf{v}_2, ..., \mathbf{v}_n\}$ satisfying $A\mathbf{v}_i = \lambda_i \mathbf{v}_i$; the λ_i are the eigenvalues of A.*

Proof not given. ■

To reinforce, the eigenvectors here, $\mathbf{v}_i$ are taken to be *unit* vectors. Let

$$M = \begin{pmatrix} \uparrow & \uparrow & & \uparrow \\ \mathbf{v}_1 & \mathbf{v}_2 & ... & \mathbf{v}_n \\ \downarrow & \downarrow & & \downarrow \end{pmatrix}$$

be the matrix with eigenvectors in each column. Then matrix multiplication gives

$$AM = (A\mathbf{v}_1, A\mathbf{v}_2, ..., A\mathbf{v}_n) = (\lambda_1\mathbf{v}_1, \lambda_2\mathbf{v}_2, ..., \lambda_n\mathbf{v}_n),$$

since each $\mathbf{v}_i$ is an eigenvector of A with eigenvalue λ_i. Further,

$$M^T AM = \begin{pmatrix} \mathbf{v}_1^T \\ \mathbf{v}_2^T \\ \vdots \\ \mathbf{v}_n^T \end{pmatrix} (\lambda_1\mathbf{v}_1, \lambda_2\mathbf{v}_2, ..., \lambda_n\mathbf{v}_n) = \left(\mathbf{v}_i^T \lambda_j \mathbf{v}_j\right)_{1\le i,j\le n}.$$

Yet

$$\mathbf{v}_i^T \lambda_j \mathbf{v}_j = \lambda_j \mathbf{v}_i^T \mathbf{v}_j = \begin{cases} \lambda_j & \text{if } i = j \\ 0 & \text{if } i \neq j. \end{cases}$$

The first case, when $i = j$, uses the fact that the vectors are of unit length, so $\mathbf{v}_i^T \mathbf{v}_i = 1$ for all i. The second case uses the fact that different $\mathbf{v}_i$ are orthogonal, i.e. $\mathbf{v}_i^T \mathbf{v}_j = 0$ if $i \neq j$. Thus $M^T AM = D$, where D is the diagonal matrix with λ_i in the ii position, $1 \le i \le n$, zero elsewhere.

Note the same argument gives

$$M^T M = \left(\mathbf{v}_i^T \mathbf{v}_j\right)_{1\le i,j\le n} = I_n.$$

Thus $M^T = M^{-1}$. Hence $M^{-1}AM = D$ or, equivalently, $A = MDM^{-1} = MDM^T$. Then for any $\mathbf{v} \in \mathbb{R}^n$ we have

$$\mathbf{v}^T A\mathbf{v} = \mathbf{v}^T (MDM^T)\mathbf{v} = (M^T\mathbf{v})^T D(M^T\mathbf{v}) = \mathbf{w}^T D\mathbf{w},$$

with $\mathbf{w} = M^T\mathbf{v}$. Therefore

$$\mathbf{v}^T A\mathbf{v} = \mathbf{w}^T D\mathbf{w} = \lambda_1 w_1^2 + \lambda_2 w_2^2 + ... + \lambda_n w_n^2.$$

If all the $\lambda_i > 0$ then $\mathbf{v}^T A\mathbf{v} > 0$ for all $\mathbf{v} \neq \mathbf{0}$, i.e. A is *positive definite.*

Conversely, if $\mathbf{v}^T A\mathbf{v} > 0$ for all $\mathbf{v} \neq \mathbf{0}$ then, given $1 \leq i \leq n$ we can choose $\mathbf{v} = \mathbf{v}_i$ the i-th eigenvector. Then

$$\begin{aligned} 0 &< \mathbf{v}_i^T A\mathbf{v}_i = \mathbf{v}_i^T \lambda_i \mathbf{v}_i \quad \text{since } \mathbf{v}_i \text{ is an eigenvector} \\ &= \lambda_i \quad \text{since } \mathbf{v}_i^T\mathbf{v}_i = 1. \end{aligned}$$

That is $\mathbf{v}^T A\mathbf{v} > 0$ for all $\mathbf{v} \neq \mathbf{0}$ implies $\lambda_i > 0$ for all $1 \leq i \leq n$.

If all $\lambda_i < 0$ then $\mathbf{v}^T A\mathbf{v} < 0$ for all $\mathbf{v} \neq \mathbf{0}$, i.e. A is *negative definite.* The converses of this can easily be shown to hold, giving

Proposition 16 *Let $A = (a_{ij})$ be a real symmetric $n \times n$ matrix.*

A is positive definite if, and only if, all eigenvalues are positive,

A is negative definite if, and only if, all eigenvalues are negative.

If some $\lambda_k > 0$ and some $\lambda_\ell < 0$ then, with the eigenvectors $\mathbf{v}_k$ and $\mathbf{v}_\ell$, we have $\mathbf{v}_k^T A\mathbf{v}_k > 0$ and $\mathbf{v}_\ell^T A\mathbf{v}_\ell < 0$, i.e. A is *indefinite.*

There is a final case where either all $\lambda_i \geq 0$ with some $\lambda_j = 0$ or all $\lambda_i \leq 0$ with some $\lambda_j = 0$. The matrix would then be nondefinite (with $\mathbf{v}_j$ the j-th eigenvector, $\mathbf{v}_j^T A\mathbf{v}_j = 0$). But note from $A = MDM^{-1}$ that $\det A = \det D = \lambda_1\lambda_2...\lambda_n$. So the present case is excluded if we assume $\det A \neq 0$.

The following was seen in Chapter 5.

Example 17 *Let $f : \mathbb{R}^3 \to \mathbb{R}$ be given by $f(\mathbf{x}) = x^2y^2 + z^2 + 2x - 4y + z$. Find the critical values of f and determine whether they are extremal values or saddle points.*

Solution The gradient vector is

$$\nabla f(\mathbf{x}) = \begin{pmatrix} 2xy^2+2 \\ 2x^2y-4 \\ 2z+1 \end{pmatrix}.$$

Solving $\nabla f(\mathbf{x}) = 0$ for critical points we have $xy^2 = -1$, $x^2y = 2$ and $z = -1/2$. Squaring the second equation gives $4 = x^4y^2 = -x^3$, using the first equation. Thus $x = -2^{2/3}$. In $x^2y = 2$ this gives $y = 2^{-1/3}$. Hence the only critical point is $\mathbf{a} = \left(-2^{2/3}, 2^{-1/3}, -2^{-1}\right)^T$.

The Hessian matrix is

$$Hf(\mathbf{x}) = \begin{pmatrix} y^2 & 4xy & 0 \\ 4xy & x^2 & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

At the critical point

$$Hf(\mathbf{a}) = \begin{pmatrix} \frac{1}{2}\sqrt[3]{2} & -4\sqrt[3]{2} & 0 \\ -4\sqrt[3]{2} & 2\sqrt[3]{2} & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

This is non-singular, in fact $\det Hf(\mathbf{a}) = -30 \times 2^{2/3}$, and so $\mathbf{a}$ is a non-degenerate critical point.

The eigenvalues of $Hf(\mathbf{a})$ are 2, $5\sqrt[3]{2}/4 + \sqrt{265}\sqrt[3]{2}/4 > 0$ and $5\sqrt[3]{2}/4 - \sqrt{265}\sqrt[3]{2}/4 < 0$ so $\mathbf{a}$ is a saddle point. The complicated nature of the eigenvalues are evidence that a simpler method is required.

7.4.1 Principal Minors

Calculating eigenvalues is difficult and we only need to know the signs of them. Can we deduce anything about the signs of the eigenvectors without calculating them?

Definition *Let $A = (a_{ij})$ be a real, symmetric, $n \times n$ matrix. For $1 \le \ell \le n$, form the $\ell \times \ell$ matrix $A_\ell = (a_{ij})_{1 \le i,j \le \ell}$. The A_ℓ are called the $\ell \times \ell$* ***principal minors*** *of A.*

The connection (if any) between the eigenvalues of A_ℓ with $\ell < n$ with the eigenvalues of $A_n = A$ is not obvious. But we have a weak connection in

Lemma 18 *If A is a real symmetric $n \times n$ matrix and A_{n-1}, the $(n-1) \times (n-1)$ principal minor of A, has $n-1$ positive eigenvalues then A has at least $n-1$ positive eigenvalues.*

Proof The assumption that A_{n-1} has $n-1$ positive eigenvalues implies A_{n-1} is positive definite.

Since A is symmetric, $\mathbb{R}^n$ has an orthonormal basis $\mathbf{v}_1, ..., \mathbf{v}_n$ of eigenvectors of A; so $A\mathbf{v}_j = \lambda_j \mathbf{v}_j$ for $1 \leq j \leq n$. If the conclusion of the lemma is false, at least two of the eigenvalues of A, say λ_1, λ_2 are < 0. Their associated eigenvalues $\mathbf{v}_1$ and $\mathbf{v}_2$ satisfy $\mathbf{v}_i^T A \mathbf{v}_i = \lambda_i \mathbf{v}_i^T \mathbf{v}_i < 0$ for $i = 1, 2$.

These $\mathbf{v}_1$ and $\mathbf{v}_2$ are also linearly independent so we can find a linear combination $\mathbf{w} = \alpha \mathbf{v}_1 + \beta \mathbf{v}_2$, $\alpha, \beta \in \mathbb{R}$ not both zero, such that $w^n = 0$, i.e.

$$\mathbf{w}^T = \left(w^1, w^2, ..., w^{n-1}, 0\right) = \left(\overline{\mathbf{w}}^T, 0\right),$$

where $\overline{\mathbf{w}} \in \mathbb{R}^{n-1}, \overline{\mathbf{w}} \neq \mathbf{0}$. Then, since $\mathbf{v}_1$ and $\mathbf{v}_2$ are orthogonal we have

$$\mathbf{v}_i^T A \mathbf{v}_j = \lambda_j \mathbf{v}_i^T \mathbf{v}_j = 0$$

for $i \neq j$. Thus

$$\begin{aligned}
\mathbf{w}^T A \mathbf{w} &= (\alpha \mathbf{v}_1 + \beta \mathbf{v}_2)^T A (\alpha \mathbf{v}_1 + \beta \mathbf{v}_2)^T \\
&= \alpha^2 \mathbf{v}_1^T A \mathbf{v}_1 + \alpha\beta \left(\mathbf{v}_2^T A \mathbf{v}_1 + \mathbf{v}_1^T A \mathbf{v}_2\right) + \beta^2 \mathbf{v}_1^T A \mathbf{v}_1 \\
&= \alpha^2 \mathbf{v}_1^T A \mathbf{v}_1 + \beta^2 \mathbf{v}_1^T A \mathbf{v}_1 < 0.
\end{aligned}$$

Yet

$$\mathbf{w}^T A \mathbf{w} = \left(\overline{\mathbf{w}}^T, 0\right) A \left(\overline{\mathbf{w}}^T, 0\right)^T = \overline{\mathbf{w}}^T A_{n-1} \overline{\mathbf{w}} > 0$$

since A_{n-1} is positive definite. This contradiction means our assumption is false and at most one eigenvalue of A can be negative. ■

The following is the promised test for positive definiteness.

Theorem 19 *If A is a symmetric $n \times n$ matrix then A is positive definite if, and only if* $\det A_\ell > 0$ *for all* $1 \leq \ell \leq n$.

Proof

$(\implies)$ Assume A is positive definite, so all eigenvalues are positive. Then $\det A = \lambda_1\lambda_2...\lambda_n > 0$.

Assume for contradiction that $\det A_\ell \le 0$ for some $1 \le \ell < n$. The matrix A_ℓ is still real symmetric and so has ℓ eigenvalues and $\det A_\ell$ is the product of these. Therefore, since $\det A_\ell \le 0$, not all of the eigenvalues of A_ℓ are positive and thus A_ℓ is not positive definite. Thus we can find a non-zero $\mathbf{w}_0 = (w_1, w_2, ..., w_\ell) \in \mathbb{R}^\ell$ for which $\mathbf{w}_0^T A_\ell \mathbf{w}_0 \le 0$. Then define $\mathbf{w} = (w_1, w_2, ..., w_\ell, 0, ..., 0) \in \mathbb{R}^n$. For this it can be checked that $\mathbf{w}^T A \mathbf{w} = \mathbf{w}_0^T A_\ell \mathbf{w}_0 \le 0$ contradicting the fact that A is positive definite. Hence $\det A_\ell > 0$ for all $1 \le \ell < n$, as required.

$(\impliedby)$ We prove the result for $n \ge 2$ by induction.

Base Case If $n = 2$ then $\det A_2 > 0$ and $\det A_1 > 0$ translate to $\det A > 0$ and $a > 0$ where

$$A = \begin{pmatrix} a & b \\ b & c \end{pmatrix}.$$

By Theorem 7 A is positive definite.

Inductive step. Assume that if B is a symmetric $(n-1)\times(n-1)$ matrix and $\det B_\ell > 0$ for all $1 \le \ell \le n-1$ then B is positive definite.

Let A be a symmetric $n \times n$ matrix with $\det A_\ell > 0$ for all $1 \le \ell \le n$, $n \ge 2$.

Then A_{n-1} is a symmetric $(n-1)\times(n-1)$ matrix. Thinking carefully about the definition of principal minor we see that $(A_{n-1})_\ell = A_\ell$ for all $1 \le \ell \le n-1$. Thus $\det(A_{n-1})_\ell = \det A_\ell > 0$ for such ℓ. Thus we have that A_{n-1} is a symmetric $(n-1)\times(n-1)$ matrix with $\det(A_{n-1})_\ell > 0$ for all $1 \le \ell \le n$. Apply the inductive hypothesis with $B = A_{n-1}$ to deduce that A_{n-1} is positive definite.

We have only used $\det A_\ell > 0$ for $\ell \le n-1$ so we still have $\det A_n > 0$, i.e. $\det A > 0$ to use.

We have A_{n-1} is positive definite and $\det A > 0$. By Lemma 18 A_{n-1} positive definite implies A has at least $n-1$ positive eigenvalues. But the product of all n eigenvalues is $\det A > 0$, so it is impossible for the n-th eigenvalue to be non-positive. Therefore all eigenvalues of A are positive and hence A is positive definite.

By induction the result holds for all $n \geq 2$. ■

So if $\det Hf(\mathbf{a})_l > 0$ for all $1 \leq \ell \leq n$ then f has a *local minimum* at the critical point $\mathbf{a}$.

Note that if M is an $n \times n$ matrix then $\det(-M) = (-1)^n \det M$. Hence

Corollary 20 *Then A is negative definite iff $-A$ is positive definite iff $\det(-A)_\ell > 0$, that is $(-1)^\ell \det A_\ell > 0$ for all $1 \leq \ell \leq n$.*

So if $(-1)^\ell \det Hf(\mathbf{a})_l > 0$ for all $1 \leq \ell \leq n$ then f has a *local maximum* at the critical point $\mathbf{a}$.

If f has neither a local minimum or maximum it is a saddle.

Example 3 continued For $f(\mathbf{x}) = x^2y + y^2z + z^2 - 2x$ the Hessian matrix at the critical point $\mathbf{a} = (1, 1, -1/2)^T$ is

$$\begin{pmatrix} 2 & 2 & 0 \\ 2 & -1 & 2 \\ 0 & 2 & 2 \end{pmatrix}.$$

Then $\det Hf(\mathbf{a})_1 = \det(2) = 2 > 0$,

$$\det Hf(\mathbf{a})_2 = \det \begin{pmatrix} 2 & 2 \\ 2 & -1 \end{pmatrix} = -6,$$

and $\det Hf(\mathbf{a})_3 = -20$. We have neither $\det Hf(\mathbf{a})_\ell > 0$ for all ℓ or $(-1)^\ell \det Hf(\mathbf{a})_\ell > 0$ for all ℓ. Hence $\mathbf{a}$ is saddle point. ■

Returning to Example 17, the function $f : \mathbb{R}^3 \to \mathbb{R}$ given by $f(\mathbf{x}) = x^2y^2 + z^2 + 2x - 4y + z$. has one critical point is $\mathbf{a} = \left(-2^{2/3}, 2^{-1/3}, -2^{-1}\right)^T$ where the Hessian matrix is

$$Hf(\mathbf{a}) = \begin{pmatrix} \frac{1}{2}\sqrt[3]{2} & -4\sqrt[3]{2} & 0 \\ -4\sqrt[3]{2} & 2\sqrt[3]{2} & 0 \\ 0 & 0 & 2 \end{pmatrix}.$$

Looking at the determinants of the principal minors

$$\det\left(\frac{1}{2}\sqrt[3]{2}\right) > 0, \;\; \det\begin{pmatrix} \frac{1}{2}\sqrt[3]{2} & -4\sqrt[3]{2} \\ -4\sqrt[3]{2} & 2\sqrt[3]{2} \end{pmatrix} < 0 \;\text{ and }\; \det Hf(\mathbf{a}) < 0.$$

This again shows that at **a** there is a saddle point. ■