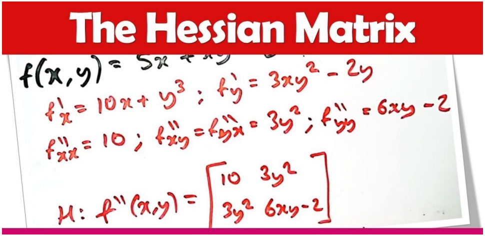
MATH 223: Multivariable Calculus



Class 20: October 28, 2022

Handouts

All Online

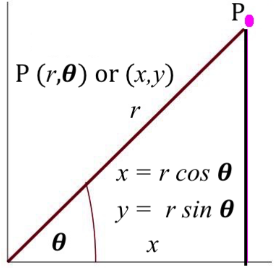
Notes on Assignment 18

Assignment 19

Ludwig Otto Hesse

ANNOUNCEMENTS

Review Polar Coordinates



Exam 2: Wednesday, November 2

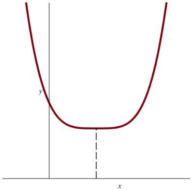
No Office Hours Today

Today:

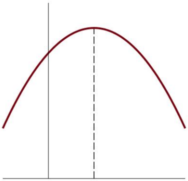
Second Derivative Criteria

Classic Case: $f : \mathbb{R}^1 \to \mathbb{R}^1$

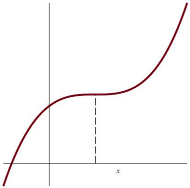
with $f'(x_0) = 0$



$f''(x_0) > 0$

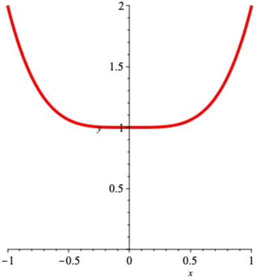


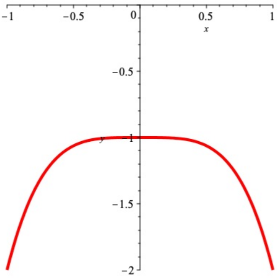
$f''(x_0) < 0$



f'' changes sign at x_0

What Can We Conclude if $f''(x_0) = 0$?





$$f(x) = x^4$$
$$f'(x) = 4x^3$$
$$f''(x) = 12x^2$$
$$f''(0) = 0$$

$$f(x) = -x^4$$
$$f'(x) = -4x^3$$
$$f''(x) = -12x^2$$
$$f''(0) = 0$$

Second Derivative Test
For Real-Valued Functions of Several Variables

Involves Second Order Partial Derivatives

Definition: If f is a twice differentiable function from $\mathbb{R}^n$ to $\mathbb{R}^1$, then the **Hessian Matrix** is the $n \times n$ matrix of second order partial derivatives of f

For example, if $f : \mathbb{R}^3 \to \mathbb{R}^1$ so $w = f(x, y, z)$, then the Hessian $\mathcal{H}$ for f is

$$\mathcal{H}(f) = \begin{pmatrix} f_{xx} & f_{xy} & f_{xz} \\ f_{yx} & f_{yy} & f_{yz} \\ f_{zx} & f_{zy} & f_{zz} \end{pmatrix}$$

Note that if f is twice continuously differentiable, then the mixed partials are equal: $f_{xy} = f_{yx}, f_{xz} = f_{zx}, f_{yz} = f_{zy}$ so the Hessian matrix is symmetric.

Otto Ludwig Hesse
April 22, 1811 – August 4, 1874

The Second Derivative Test
for real-valued functions of several variables
replaces the condition $f''(c)$ being positive
with the Hessian matrix being **positive definite**.
It similarly uses the **negative definite** character of the Hessian matrix in place of the negativity of the second derivative.

Definition A *Positive Definite Matrix* is an n by n symmetric matrix A such that $\mathbf{x} \cdot (A\mathbf{x}) > 0$ for all nonzero vectors $\mathbf{x}$ in $\mathbb{R}^n$.

You will often see the equivalent condition $\mathbf{x}^T A\mathbf{x} > 0$ for $\mathbf{x} \neq \mathbf{0}$ where $\mathbf{x}^T$ is the transpose of $\mathbf{x}$.

If the strict inequality sign $>$ is replaced by the weaker $\geq$, then the matrix is called *positive semi-definite*.

We define *negative definite* and *negative semi-definite* in an analogous manner, using $<$ and $\leq$, respectively.

Example: Let $A = \begin{pmatrix} 10 & 4 \\ 4 & 2 \end{pmatrix}$.

With $\mathbf{x} = \begin{pmatrix} x \\ y \end{pmatrix}$, we have

$$\begin{aligned}(x, y) \cdot (10x + 4y, 4x + 2y) &= 10x^2 + 4xy + 4xy + 2y^2 \\ &= 10x^2 + 8xy + 2y^2 \\ &= (9x^2 + 6xy + y^2) + (x^2 + 2xy + y^2) \\ &= (3x + y)^2 + (x + y)^2\end{aligned}$$

which is the sum of non-negative numbers and hence always greater than or equal to 0, but is positive unless both x and y are 0. Hence A is a positive definite matrix.

A Matrix Which is Not Positive Definite

$$A = \begin{pmatrix} 2 & 4 \\ 4 & 5 \end{pmatrix}$$

With $x = -2, y = 1$, we have $\mathbf{x} \cdot (A\mathbf{x}) = -3$ so A is not positive definite. With $x = 2, y = 1$, we have $\mathbf{x} \cdot (A\mathbf{x}) = 29$ so A is not negative definite.

An alternative, equivalent definition of a symmetric matrix being positive definite is that all its eigenvalues are positive.

Theorem

All real eigenvalues of a positive definite matrix are positive.

Proof: Let A be an $n \times n$ be a positive definite matrix with real eigenvalue λ.

If $\lambda = 0$, then there is a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = 0\mathbf{x} = \mathbf{0}$. But then, $\mathbf{x} \cdot (A\mathbf{x}) = 0$ so A would not be positive definite. □

If $\lambda < 0$, then there is a nonzero vector $\mathbf{x}$ such that $A\mathbf{x} = \lambda\mathbf{x}$ in which case

$$\mathbf{x} \cdot (A\mathbf{x}) = \mathbf{x} \cdot (\lambda\mathbf{x}) = \lambda|\mathbf{x}|^2 < 0$$

so again A is not positive definite. □

Not only is the converse of this theorem true (all eigenvalues positive implies positive definiteness), but the eigenvalues of a symmetric matrix are always real.

Theorem

If all the eigenvalues of a symmetric matrix A are positive, then A is positive definite.

Proof: We make use of a result from linear algebra: A symmetric matrix is diagonalizable by an orthogonal matrix; that is, there is an orthogonal matrix Q such that $Q^T = Q^{-1}$ with $Q^T AQ = D$, where D is a diagonal matrix whose main diagonal entries are the eigenvalues of A:

$$D = \begin{pmatrix} \lambda_1 & 0 & 0 & \dots & 0 \\ 0 & \lambda_2 & 0 & \dots & 0 \\ 0 & 0 & \lambda_3 & \dots & 0 \\ \dots & & & & \\ 0 & 0 & 0 & \dots & \lambda_n \end{pmatrix}$$

Let $\mathbf{x}$ be any nonzero vector and set $\mathbf{y} = Q^T\mathbf{x}$ so that $\mathbf{y}^T = \mathbf{x}^T Q$. Then

$$\mathbf{x}^T A\mathbf{x} = \mathbf{x}^T(QDQ^T)\mathbf{x} = (\mathbf{x}^T(Q)D(Q^T\mathbf{x}^T) = \mathbf{y}^T D\mathbf{y}$$

but

$$\mathbf{y}^T D\mathbf{y} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2 = \sum_{i=1}^{n} \lambda_i y_i^2$$

where $\mathbf{y}^T = (y_1, y_2, \ldots, y_n)$. Since $\mathbf{x}$ is a nonzero vector and Q is invertible, at least one y_i is nonzero. Hence

$$\mathbf{x}^T A\mathbf{x} = \lambda_1 y_1^2 + \lambda_2 y_2^2 + \ldots + \lambda_n y_n^2 = \sum_{i=1}^{n} \lambda_i y_i^2$$

is the sum of non-negative terms, at least one of which is positive, so it is positive. Thus A is positive definite. $\square$.

Now, we'll take a look at the occurrence of positive definite matrices in testing critical points for local extreme properties. The setting is a real -valued function $f : \mathbb{R}^n \to \mathbb{R}^1$ which has continuous partial derivatives of third order in an open set U containing a vector $\mathbf{x_0}$.

The Second-Order Taylor Theorem asserts that

$$f(\mathbf{x_0} + \mathbf{h}) = f(\mathbf{x_0}) + \nabla f(\mathbf{x_0}) \cdot \mathbf{h} + \frac{1}{2}\mathbf{h}^T \mathcal{H}(\mathbf{x_0})\mathbf{h} + R_2(\mathbf{x_0}, \mathbf{h})$$

and

$$\lim_{\mathbf{h} \to 0} \frac{R_2(\mathbf{x_0}, \mathbf{h})}{|\mathbf{h}|} = 0$$

(See Text for Proof)

Theorem

Second Derivative Test for Local Extrema.

Suppose $f : \mathbb{R}^n \to \mathbb{R}^1$ *has continuous third order partial derivatives on a neighborhood of* $\mathbf{x_0}$ *which is a critical point of* f *.*

IF the Hessian $\mathcal{H}$ *evaluated at* $\mathbf{x_0}$ *is positive definite, then* f *has a relative minimum at* $\mathbf{x_0}$*.*

If the Hessian is negative definite, then there is a relative maximum at the critical point.

Here is the idea of the proof: Since $\mathbf{x_0}$ is a critical point, $\nabla f(\mathbf{x_0}) = \mathbf{0}$ and by Taylor's Theorem

$$f(\mathbf{x_0} + \mathbf{h}) = f(\mathbf{x_0}) + \frac{1}{2}\mathbf{h}^T \mathcal{H}(\mathbf{x_0})\mathbf{h} + R_2(\mathbf{x_0}, \mathbf{h})$$

where the remainder term is negligible when $\mathbf{h}$ is very small. Thus

$$f(\mathbf{x_0} + \mathbf{h}) \approx f(\mathbf{x_0}) + \frac{1}{2}\mathbf{h}^T \mathcal{H}(\mathbf{x_0})\mathbf{h}.$$

If the Hessian is positive definite, then the second term is positive for $\mathbf{h} \neq \mathbf{0}$ so $f(\mathbf{x_0} + \mathbf{h}) > f(\mathbf{x_0})$ when $\mathbf{h}$ is sufficiently small, making $\mathbf{x_0}$ the location of a relative minimum.
We will leave a formal proof and dealing with the negative definite case for the exercises.

∎

Example: Our Temperature Function

$$T(x, y) = 2x^2 + 4y^2 + 2x + 1$$

Here $T_x(x, y) = 4x + 2$ and $T_y(x, y) = 8y$.

Thus, the Hessian Matrix $\mathcal{H}$ is

$$\mathcal{H} = \begin{pmatrix} 4 & 0 \\ 0 & 8 \end{pmatrix}$$

whose eigenvalues are 4 and 8.

Both are positive so T has a minimum wherever the gradient is 0; that is, at (-1/2,0).

Example: $T(x, y) = x^2 - y^2$

$\nabla T(x, y) = (2x, -2y)$ so $\nabla T = \vec{0}$ at $(0, 0)$

Thus, the Hessian Matrix $\mathcal{H}$ is

$$\mathcal{H} = \begin{pmatrix} 2 & 0 \\ 0 & -2 \end{pmatrix}$$

whose eigenvalues are 2 and -2.
Thus it is neither positive definite nor negative definite.

$\mathbf{x} \cdot \mathcal{H}\mathbf{x}$ can be positive ($\mathbf{x} = (1, 0)$) or negative ($\mathbf{x} = (0,1)$) so there is a **saddle point** at any point where ∇T is $\vec{0}$.

Example: $f(x, y) = x^3 - y^3 - 2xy$

Here $\nabla f = (3x^2 - 2y, -3y^2 - 2x)$

$\nabla f = \vec{0}$ when
$3x^2 = 2y$ and $3y^2 = -2x$

The first equation gives $9x^4 = 4y^2$ and the second yields $y^2 = -\frac{2}{3}x$
Thus $9x^4 = 4(-\frac{2}{3}x) = -\frac{8}{3}x$
so $9x^4 = -\frac{8}{3}x$ or $27x^4 + 8x = 0$; Hence $x(27x^3 + 8) = 0$
This gives two solutions: $x = 0, y = 0$ and $x = -\frac{2}{3}, y = \frac{2}{3}$

Two Critical Points: (0,0) and (-2/3, 2/3)

Example: $f(x,y) = x^3 - y^3 - 2xy$

$\nabla f = (3x^2 - 2y, -3y^2 - 2x)$

Two Critical Points: (0,0) and (-2/3, 2/3)

The Hessian Matrix is

$$\mathcal{H} = \begin{pmatrix} 6x & -2 \\ -2 & -6y \end{pmatrix}$$

At (-2/3,2/3), Hessian is

$$\mathcal{H} = \begin{pmatrix} -4 & -2 \\ -2 & -4 \end{pmatrix}$$

whose eigenvalues are -2 and -6, both negative.
Thus there is a relative maximum at (-2/3,2/3),/

Example: $f(x,y) = x^3 - y^3 - 2xy$
$\nabla f = (3x^2 - 2y, -3y^2 - 2x)$
Two Critical Points: (0,0) and (-2/3, 2/3)
The Hessian Matrix is

$$\mathcal{H} = \begin{pmatrix} 6x & -2 \\ -2 & -6y \end{pmatrix}$$

At (0,0), Hessian is

$$\mathcal{H} = \begin{pmatrix} 0 & -2 \\ -2 & 0 \end{pmatrix}$$

whose eigenvalues are -2 and +2.

.

Thus there is a saddle point at (0,0),

More About Saddle Points

"Relative Maximum in One Direction, but Relative Minimum in Another Direction"

How Do We Find These Directions?

Look at the Eigenvectors!

Take our example $f(x,y) = x^3 - y^3 - 2xy$ at the origin.

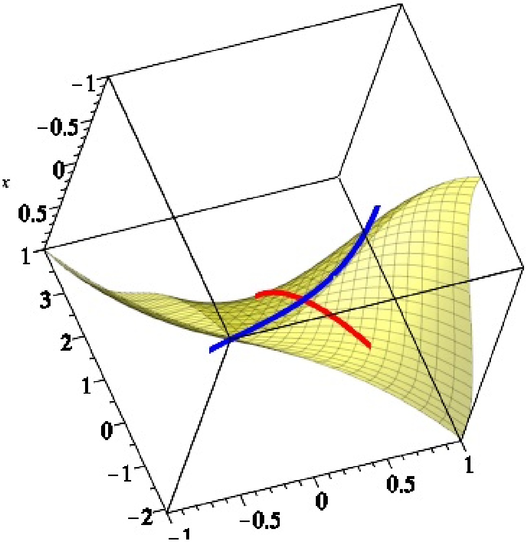
The eigenvalue -2 has eigenvector of the form (1,1)

Consider $f(x,x) = x^3 - x^3 - 2xx = -2x^2$ has relative maximum at $x = 0$

The eigenvalue $+2$ has eigenvector of the form (1,-1).

Consider $f(x,-x) = x^3 + x^3 + 2xx = 2x^3 + 2x^2 = 2x^2(1+x)$ has relative minimum at $x = 0$

Graph of $f(x,y) = x^3 - y^3 - 2xy$



Next Time

Alternative Coordinate Systems for 3-Space

Rectangular
Cylindrical
Spherical