Lecture 5-6. Taylor expansions for many variables.

Aim: - To learn how to write out the Taylor expansions for many variables.
- To classify the stationary points for the case of two variables

Reminder: 1 variable $y = f(x)$ near $x = x_0$

What is behind the Taylor expansions?

Want to approximate f near x_0 by a polynomial of order n:

$$(*)\quad f(x) = A_1 + A_2(x-x_0) + A_3(x-x_0)^2 + \dots + A_{n+1}(x-x_0)^n + O((x-x_0)^{n+1})$$

up to the error of order $(x-x_0)^{n+1}$

What are the coefficients A_k?

Is it possible to do?

Heuristics: If $(*)$ is true, then taking $x = x_0$,

$$A_1 = f(x_0)$$

Differentiating $(*)$ by x

$$f'(x) = A_2 + 2A_3(x-x_0) + \dots + nA_{n+1}(x-x_0)^{n-1} + O((x-x_0)^n)$$

taking $x = x_0 \Rightarrow$

$$A_2 = f'(x_0)$$

Repeating this, we see

$$2A_3 = f''(x_0) \Rightarrow A_3 = \frac{f''(x_0)}{2!}$$

$$1 \cdot 2 \cdots (k-1)k\, A_{k+1} = f^{(k)}(x_0)$$

$$\Rightarrow A_{k+1} = \frac{f^{(k)}(x_0)}{k!}$$

$\Rightarrow$ if $(*)$ is true then, necessarily,

$$f(x) = f(x_0) + \frac{f'(x_0)}{1!}(x-x_0) + \frac{f''(x_0)}{2!}(x-x_0)^2 + \dots + \frac{f^{(n)}(x_0)}{n!}(x-x_0)^n + O((x-x_0)^{n+1})$$

NB: It is not a rigorous proof of $(*)$! Just a heuristic derivation! Can be justified for smooth functions using mean value theorem.

Return to many variables:

For simplicity, many = 2

$z = f(x, y)$ near $(x, y) = (x_0, y_0)$

$\delta x := x - x_0$, $\delta y = y - y_0$

Analogously to 1D case, want to approximate f by polynomials of order n.

What is the main difference?

1D: only one monomial of order n: $(\delta x)^n$

2D: $(n+1)$ – monomials of order n: $(\delta x)^n, (\delta x)^{n-1}\delta y, \dots, (\delta x)(\delta y)^{n-1}, (\delta y)^n$

For instance, for $n=2$, we have 3 monomials of order 2:

$(x-x_0)^2$, $(x-x_0)(y-y_0)$, $(y-y_0)^2$

Thus, we need to write

(**) $f(x,y) = A_0 + B_1(x-x_0) + B_2(y-y_0) + [C_1(x-x_0)^2 + C_2(x-x_0)(y-y_0) + C_3(y-y_0)^2] + \text{H.O.T (cubic and higher)}$

0-order; linear; quadratic

Need to find A_0, B_i, C_i ?

Again heuristics:

take $x = x_0$, $y = y_0$ in (**) $\Rightarrow$ $A_0 = f(x_0, y_0)$

Diff. (**) with respect to x:

(1) $f'_x(x,y) = B_1 + 2C_1(x-x_0) + C_2(y-y_0) + \text{H.O.T}$

$\Rightarrow$ $B_1 = f'_x(x_0, y_0)$ – we already know this

Diff. (**) by y

(2) $f'_y(x,y) = B_2 + C_2(x-x_0) + 2C_3(y-y_0) + \text{H.O.T}$

$B_2 = f'_y(x_0, y_0)$ – also known

$\delta f = f'_x \delta x + f'_y \delta y$

Now, diff. (1) by x

$$f''_{xx}(x,y) = 2C_1 + H.O.T.$$

$$\Rightarrow \quad C_1 = \frac{f''_{xx}(x_0,y_0)}{2!}$$ – new information!

diff. (2) by y

$$f''_{yy}(x,y) = C_3 + H.O.T$$

$$\Rightarrow \quad C_3 = \frac{f''_{yy}(x_0,y_0)}{2!}$$ – new information!

it remains to find C_2.

Two ways:

1) Diff. (1) by y.

$$f''_{xy}(x,y) = C_2 + H.O.T.$$

$$\Rightarrow \quad C_2 = f''_{xy}(x_0,y_0)$$

2) Diff. (2) by x:

$$f''_{yx}(x,y) = C_2 + H.O.T.$$

$$\Rightarrow \quad C_2 = f''_{yx}(x_0,y_0)$$

Conclusion: we expect

$$f''_{xy}(x,y) = f''_{yx}(x,y)$$

In other words: differential operators $\frac{\partial}{\partial x}$ and $\frac{\partial}{\partial y}$ commute

$$\frac{\partial}{\partial x}\cdot\frac{\partial}{\partial y} f = \frac{\partial}{\partial y}\cdot\frac{\partial}{\partial x} f$$ – NB!!!

The order of taking partial derivatives is not important!

Again, not a rigorous proof, just heuristics.

Theorem: indeed commute for smooth f.

NB: may be not true if f is not smooth (!)

Final result:

$$f(x,y) = \underbrace{f(x_0,y_0)}_{\text{0 order}} + \underbrace{f'_x(x_0,y_0)(x-x_0) + f'_y(x_0,y_0)(y-y_0)}_{\text{linear}} + \frac{1}{2!}\Big[\underbrace{f''_{xx}(x_0,y_0)(x-x_0)^2 + 2f''_{xy}(x_0,y_0)(x-x_0)(y-y_0) + f''_{yy}(x_0,y_0)(y-y_0)^2}_{\text{quadratic}}\Big] + H.O.T$$

Taylor expansions for 2 variables

Equivalent form:

$$\delta f = f'_x \delta x + f'_y \delta y + \frac{1}{2!}\left[f''_{xx}\delta x^2 + 2f''_{xy}\delta x\delta y + f''_y \delta y\right] + H.O.T$$

Vector form: — matrix!!

$$\delta f = \nabla f \cdot \begin{pmatrix}\delta x\\ \delta y\end{pmatrix} + \frac{1}{2!}\begin{pmatrix}f''_{xx} & f''_{xy}\\ f''_{xy} & f''_{yy}\end{pmatrix}\begin{pmatrix}\delta x\\ \delta y\end{pmatrix}\cdot\begin{pmatrix}\delta x\\ \delta y\end{pmatrix} + H.O.T$$

dot product

Indeed,

$$\begin{pmatrix}f''_{xx} & f''_{xy}\\ f''_{xy} & f''_{yy}\end{pmatrix}\begin{pmatrix}\delta x\\ \delta y\end{pmatrix} = \begin{pmatrix}f''_{xx}\delta x + f''_{xy}\delta y\\ f''_{xy}\delta x + f''_{yy}\delta y\end{pmatrix}$$

$$\begin{pmatrix}f''_{xx} & f''_{xy}\\ f''_{xy} & f''_{yy}\end{pmatrix}\begin{pmatrix}\delta x\\ \delta y\end{pmatrix}\cdot\begin{pmatrix}\delta x\\ \delta y\end{pmatrix} = (f''_{xx}\delta x + f''_{xy}\delta y)\delta x + (f''_{xy}\delta x + f''_{yy}\delta y)\delta y = f''_{xx}\delta x^2 + 2f''_{xy}\delta x\delta y + f''_{yy}\delta y^2$$

The matrix has a special name: Hessian of f.

$$H(f) := \begin{pmatrix}f''_{xx} & f''_{xy}\\ f''_{xy} & f''_{yy}\end{pmatrix}$$

Thus,

$$\delta f = \nabla f|_{(x_0,y_0)} \cdot \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} + \frac{1}{2!} H(f)|_{(x_0,y_0)} \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} \cdot \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} + \text{H.O.T.}$$

= Taylor expansions up to order 2 inclusively.

NB: 1D case = f' and f'' are scalars

many variables:

First derivative is a vector (gradient)

$$\nabla f = (f'_x, f'_y)$$

Second derivative is a matrix (Hessian)

$$H(f) = \begin{pmatrix} f''_{xx}, f''_{xy} \\ f''_{xy}, f''_{yy} \end{pmatrix}$$

Example: $f(x,y) = e^{x+y} + \cos(x-y)$

$$f'_x = e^{x+y} - \sin(x-y)$$
$$f'_y = e^{x+y} + \sin(x-y)$$

$$\nabla f = \underbrace{\left(e^{x+y} - \sin(x-y),\ e^{x+y} + \sin(x-y) \right)}_{\text{vector}}$$

$$f''_{xx} = e^{x+y} - \cos(x-y), \quad f''_{yy} = e^{x+y} - \cos(x-y)$$
$$f''_{xy} = e^{x+y} + \cos(x-y), \quad f''_{yx} = e^{x+y} + \cos(x-y)$$

As predicted

$$f''_{xy} = f''_{yx}$$

Hessian:

$$H(f)=\begin{pmatrix} e^{x+y}-\cos(x-y), & e^{x+y}+\cos(x-y) \\ e^{x+y}+\cos(x-y), & e^{x+y}-\cos(x-y) \end{pmatrix}$$

$= 2\times 2$ matrix

Find Taylor expansions up to order 2 inclusively near $(x,y)=(\pi,-\pi)$.

Solution

$$\nabla f\big|_{(\pi,-\pi)} = (1,1)$$

$$H(f)\big|_{(\pi,-\pi)} = \begin{pmatrix} 0, & 2 \\ 2, & 0 \end{pmatrix}$$

$$f(\pi,-\pi)=2$$

$$f(x,y) = 2 + (x-\pi)+(y+\pi)+\frac{1}{2}\left[2\cdot 2\,(x-\pi)(y+\pi)\right] =$$

$$= 2+(x-\pi)+(y+\pi)+2(x-\pi)(y+\pi)+H.O.T$$

Extension to 3-variables: completely analogous

$$f = f(x,y,z)$$

$$\nabla f = (f'_x, f'_y, f'_z)$$ – 3-vector (gradient)

$$H(f)=\begin{pmatrix} f''_{xx}, & f''_{xy}, & f''_{xz} \\ f''_{xy}, & f''_{yy}, & f''_{yz} \\ f''_{xz}, & f''_{yz}, & f''_{zz} \end{pmatrix}$$ – Hessian (3×3-matrix)

Taylor expansions:

$$\delta f = \underbrace{\nabla f\cdot\begin{pmatrix}\delta x\\ \delta y\\ \delta z\end{pmatrix}}_{\text{linear}} + \underbrace{\frac{1}{2}H(f)\begin{pmatrix}\delta x\\ \delta y\\ \delta z\end{pmatrix}\cdot\begin{pmatrix}\delta x\\ \delta y\\ \delta z\end{pmatrix}}_{\text{quadratic deviation from linear}} + H.O.T$$

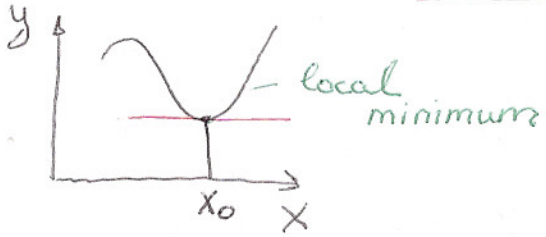
OR

$$f(x,y) = f(x_0,y_0,z_0) + f'_x(x_0,y_0,z_0)(x-x_0) + f'_y(\cdot)(y-y_0)+$$
$$+ f'_z(\cdot)(z-z_0) + \frac{1}{2!}\left(f''_{xx}(\cdot)(x-x_0)^2 + 2f''_{xy}(\cdot)(x-x_0)(y-y_0)+\right.$$
$$+ 2f''_{xz}(\cdot)(x-x_0)(z-z_0) + 2f''_{yz}(\cdot)(y-y_0)+(z-z_0)+$$
$$\left.+ f''_{yy}(\cdot)(y-y_0)^2 + f''_{zz}(\cdot)(z-z_0)^2\right] + H.O.T$$

Application: Local maximum, minimum or saddle points

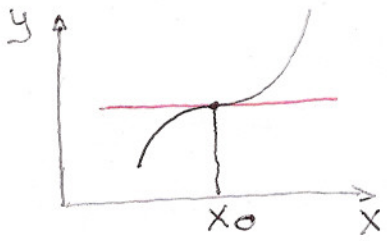
Reminder: 1D $\quad y = f(x)$

I



II

III



inflection point

Questions: a) How to find them?
b) How to distinguish between cases I, II and III?

Local analysis (in a small neighbourhood of x_0). Start with 1st order approximation:

$$\delta f = f'(x_0)\,\delta x + \text{H.O.T} \qquad (*)$$

For max/min δf must be sign definite (negative/positive)

But if $f'(x_0) \neq 0$ $\Rightarrow$ the term $f'(x_0)\delta x$ is dominating for small δx and NOT sign-definite

$\Rightarrow$ **First order test**

critical ~ stationary points.

x_0 - point of max/min $\Longrightarrow$ $f'(x_0) = 0$

NB: Necessary, but not sufficient ($y = x^3$ near 0)

Second order test is necessary

Let $f'(x_0) = 0 \Rightarrow$

$$\delta f = \frac{1}{2} f''(x_0)\,\delta x^2 + \text{H.O.T.}$$

sign-definite if $f''(x_0) \neq 0$

$\Rightarrow$ $f''(x_0) > 0$ - positive $\Rightarrow$ minimum
$f''(x_0) < 0$ - negative $\Rightarrow$ maximum

$f''(x_0) = 0$ - inconclusive $\Rightarrow$ H.O.T are necessary!

e.g. $f'''(x_0) \neq 0 \Rightarrow$ not sign-def $\Rightarrow$ inflection

Aim: To obtain something similar for many variables.

$$z = f(x, y)$$

"many" = 2 (as usual!)

Start with 1st order local analysis (near (x_0, y_0))

$$\delta f = f'_x\big|_{x_0,y_0}\,\delta x + f'_y\big|_{x_0,y_0}\,\delta y + H.O.T$$

when δf can be sign-definite (positive or negative)

if $\nabla f\big|_{x_0,y_0} \neq 0$ $\Rightarrow$ Linear part is dominating

$$\delta f \approx f'_x\,\delta x + f'_y\,\delta y$$

(linear)

But linear function cannot be sign-def!!!

Indeed if $\nabla f \cdot \begin{pmatrix}\delta x\\ \delta y\end{pmatrix} > 0 \Rightarrow \nabla f \cdot \begin{pmatrix}-\delta x\\ -\delta y\end{pmatrix} < 0$ $\forall \delta x, \delta y$

$\Rightarrow$ necessary condition = (1st order test)

(x_0, y_0) – max/min $\Rightarrow$ $\nabla f\big|_{x_0,y_0} = 0$

How to find stationary (extremal) point? stationary points!!

Need to solve 2 simultaneous equations!

$$\begin{cases} f'_x(x, y) = 0 \\ f'_y(x, y) = 0 \end{cases}$$

In contrast to diff., solving such systems is non-trivial task!!!

Only in very particular cases explicit solution can be found $\Rightarrow$ numerics is necessary!

Example:

$$f = x + y + \frac{1}{x} + \frac{1}{y}$$

$$f'_x = 1 - \frac{1}{x^2}, \quad f'_y = 1 - \frac{1}{y^2}$$

System on extremals

$$\begin{cases} 1 - \frac{1}{x^2} = 0 \\ 1 - \frac{1}{y^2} = 0 \end{cases} \Rightarrow x = \pm 1,\ y = \pm 1$$

$\Rightarrow$ (4) stationary points: $(x_0, y_0) = (1, 1), (-1, 1), (1, -1), (-1, -1)$

Types of stationary points: 2nd order test is required!

$$\delta f = \frac{1}{2} H(f)\big|_{x_0,y_0} \begin{pmatrix}\delta x\\ \delta y\end{pmatrix}\cdot\begin{pmatrix}\delta x\\ \delta y\end{pmatrix} + \text{H.O.T.}$$

$$H(f) = \begin{pmatrix} f''_{xx}, & f''_{xy}\\ f''_{xy}, & f''_{yy}\end{pmatrix} \text{ - Hessian matrix}$$

Need to know when the quadratic form

$$a\delta x^2 + 2b\delta x\cdot\delta y + c\delta y^2 = \begin{pmatrix} a, & b\\ b, & c\end{pmatrix}\begin{pmatrix}\delta x\\ \delta y\end{pmatrix}\cdot\begin{pmatrix}\delta x\\ \delta y\end{pmatrix}$$

is sign-def.

$$H(f)\big|_{x_0,y_0} = \begin{pmatrix} a, & b\\ b, & c\end{pmatrix}$$

Excursion to Linear Algebra:

Let $\vec{x} = \begin{pmatrix}x_1\\ x_2\end{pmatrix} = x_1 e_1 + x_2 e_2 \in \mathbb{R}^2$, e_1, e_2 - coordinate vectors

$A := \begin{pmatrix} a, & b\\ b, & c\end{pmatrix}$ - symmetric matrix

$$A\vec{x}\cdot\vec{x} = \vec{x}^t A\vec{x} = a x_1^2 + 2b x_1 x_2 + c x_2^2$$

Th.: Any symmetric matrix $(n\times n)$ is diagonalizable has real eigenvalues. The corresponding eigenvectors are ortogonal.

(will be proved in Linear Algebra)

In our case: $n=2$ Let λ_1, λ_2 be eigenvalues of A

$$\det\begin{pmatrix} a-\lambda, & b\\ b, & c-\lambda\end{pmatrix} = 0$$

$$\lambda^2 - (a+c)\lambda + [ac - b^2] = 0$$

$D = (a+c)^2 - 4[ac-b^2] = (a-c)^2 + 4b^2 \geq 0$

indeed, real eigenvalues.

$A\vec{p} = \lambda_1\vec{p}$; $A\vec{q} = \lambda_2\vec{q}$

$\vec{p}, \vec{q}$ - eigenvectors

$\vec{p}\cdot\vec{q} = 0$ - orthogonality, condition

normalized as $|\vec{p}| = |\vec{q}| = 1$

if $\lambda_1 \neq \lambda_2$

$$\lambda_1\vec{p}\cdot\vec{q} = A\vec{p}\cdot\vec{q} = \vec{p}\cdot A\vec{q} = \lambda_2\vec{p}\cdot\vec{q}$$

(A symmetric)

Write out the vector X in the $\vec{p}, \vec{q}$ -basis

$$X = x_1 e_1 + x_2 e_2 = x_1' \vec{p} + x_2' \vec{q}$$

$$\vec{p} = \begin{pmatrix} p_1 \\ p_2 \end{pmatrix}, \quad \vec{q} = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}$$

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} p_1 & q_1 \\ p_2 & q_2 \end{pmatrix} \begin{pmatrix} x_1' \\ x_2' \end{pmatrix}$$

orthogal matrix

in 2D - just rotation.

$$A\vec{x}.\vec{x} = A(x_1'\vec{p} + x_2'\vec{q}).(x_1'\vec{p} + x_2'\vec{q}) = (A\vec{p}.\vec{p})\,x_1'^2 + (2A\vec{p}.\vec{q})\,x_1'x_2' + (A\vec{q}.\vec{q})\,x_2'^2 = \lambda_1 (x_1')^2 + \lambda_2 (x_2')^2$$

($A\vec{p}.\vec{p} = \lambda_1$, $A\vec{p}.\vec{q} = 0$, $A\vec{q}.\vec{q} = \lambda_2$)

diagonal form: $x_1' x_2'$ disappears!

Def: Directions $\vec{p}$ and $\vec{q}$ are called principal axis of the quadratic form

Conclusion: every quadratic form can be diagonalized by the appropriate rotation (orthogonal transform)

$\lambda_1 (x_1')^2 + \lambda_2 (x_2')^2$ - true in any dimension!

just will be

$$\lambda_1 (x_1')^2 + \dots + \lambda_n (x_n')^2$$

λ_i - eigenvectors of A.

End of Linear Algebra excursion.

Important for us: Quadratic form is positive / negative $\Leftrightarrow$ All eigenvalues are positive / negative

$$\delta f = H(f)\big|_{(x_0, y_0)} \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} . \begin{pmatrix} \delta x \\ \delta y \end{pmatrix} + H.O.T.$$

Second order test:

(I) if All eigenvalues of $H(f)$ are positive

$\Rightarrow$ $\delta f > 0$ $\Rightarrow$ minimum

(II) if All eigenvalues of $H(f)$ are negative

$\Rightarrow$ $\delta f < 0$ $\Rightarrow$ maximum

III if there are both positive and negative eigenvalues
$\Rightarrow \delta f$ - is not sign-def $\Rightarrow$ saddle
new object !!

IV Zero eigenvalues $\Rightarrow$ inconclusive
higher order terms are required !

Works in Any dimension !

$n=2$ is simple. Indeed,

$$\lambda_1 \lambda_2 = \det H(f) = f''_{xx} \cdot f''_{yy} - \left(f''_{xy}\right)^2$$

$$\lambda_1 + \lambda_2 = T_2 H(f) = f''_{xx} + f''_{yy}$$

Case I : ~~maximum~~ minimum

$$\det H > 0, \quad T_2 H > 0$$

$\delta f = \lambda_1 (\delta x')^2 + \lambda_2 (\delta y')^2$ in principal axis

looks like $z = x^2 + y^2$
elliptic paraboloid

Case II : maximum

$$\det H > 0, \quad T_2 H < 0$$

looks like $z = -x^2 - y^2$
elliptic paraboloid

Case III : saddle

$$\det H < 0$$

(max) in one principal direction and (min) in the other

looks like $z = x^2 - y^2$
hyperbolic paraboloid

Case IV :

$$\det H = 0$$

- quadratic form vanishes in one of principal directions $\Rightarrow$
3rd order test is required.

Example:

$$z = x + y + \frac{1}{x} + \frac{1}{y}$$

Critical points

$(1,1), (1,-1), (-1,1), (-1,-1)$

Let us analyze them

$$H = \begin{pmatrix} 2/x^3 & 0 \\ 0 & 2/y^3 \end{pmatrix}$$

Already diagonal in that example

$$H|_{(1,1)} = \begin{pmatrix} 2 & 0 \\ 0 & 2 \end{pmatrix}$$ - local minimum

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

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

$$H|_{(-1,-1)} = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix}$$ - local maximum

One more example:

$$z = x^2 + xy + y^2$$

Already quadratic form at $(0,0)$

$$\nabla f|_{(0,0)} = 0$$

$$H(f) = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$$

$\det H = 2 > 0, \; T_2 H = 4 > 0$

0 is local (global) minimum

surface is elliptic paraboloid