Partial derivatives.

(a) Maxima & minima.
using second derivative test.

EE1203: Vector Calculus

Aneesh Sobhanan
Department of Electrical Engineering
IIT Hyderabad, India

20th January, 2025

భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Maxima & minima.

* Optimization problems.

⇝ Find min/max of a function in 2 variables $f(x,y)$

* At a local min or max $f_x = 0$ and $f_y = 0$.

* From Approx formula; $\Delta z \approx f_x \Delta x + f_y \Delta y$.

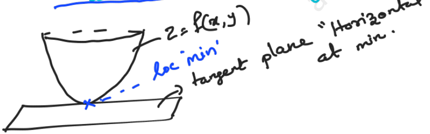
$\Rightarrow$ when $f_x = f_y = 0$ $\Delta z = 0$

$\Rightarrow$ The function does not change anymore

$\Rightarrow$ local max or local min. (need not be always true)

$\Rightarrow$ Or it also means the tangent plane is horizontal to the graph of the function $z = f(x,y)$

Ex:



* Definition: (a, b) is a critical point of $f(x,y)$ if $f_x(a,b) = 0$ and $f_y(a,b) = 0$.

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

$f_x = 2x - 2y + 2 = 0$

$f_y = -2x + 6y - 2 = 0$

$\Rightarrow \quad 4y = 0 \quad \Rightarrow y = 0.$

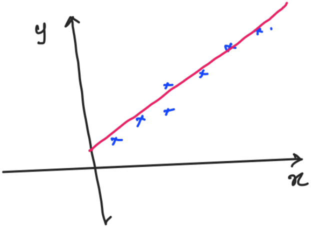
$\Rightarrow \quad 2x = -2 \quad x = -1$

$\Rightarrow$ Critical point is, $(-1, 0)$

Critical points can be: (a) Local min
(b) Local max
(c) Saddle point.

One minimization problem / optimization problem :

* Least square interpolation :



Question:
Given experimental data
$(x_1, y_1, x_2, y_2, \dots x_n, y_n)$
find best fit line
$\underline{y = ax + b}$.

Here the unknown are : a & b.

idea : find function $f(a,b)$ such that it gives minimum error between the best fit line to the measured data points.

Why least mean square errors ? (MSE)

It penalize larger error compared to smaller ones. which helps the model to optimize those outlayer points

* $\hat{y}_i \rightarrow$ measured.

$y_i \rightarrow$ predicted / fit

$$\boxed{\frac{1}{n}\sum(\hat{y}_i - y_i)^2} \rightarrow \text{MSE}$$

In terms of convex optimization problems:

* squared error function is differentiable and convex.

* This convexity ensures that there is a unique minimum MSE making optimization straight forward using methods like gradient descent.

* Convex nature guarantees that the optimization process will converge to global minimum.

Find "optimum" a & b. $\Rightarrow$ which minimize total mean square error for each data points.

$$\Rightarrow \quad \underbrace{J(a,b)}_{\text{minimize this function.}} = \frac{1}{n}\sum_{i=1}^{n}\left(y_i - (ax_i + b)\right)^2 \quad \checkmark$$

" Cost function " for linear regression!

Use idea of critical points

$$\frac{\partial J}{\partial a} = 0 \quad \& \quad \frac{\partial J}{\partial b} = 0.$$

$$\Rightarrow \quad \frac{\partial J}{\partial a} = \frac{1}{\cancel{2}n}\sum_{i=1}^{n} \cancel{2}\left[y_i - (ax_i + b)\right](-x_i) = 0.$$

$$\Rightarrow \quad \frac{\partial J}{\partial b} = \frac{1}{\cancel{2}n}\sum_{i=1}^{n} \cancel{2}\left[y_i - (ax_i + b)\right](-1) = 0.$$

$$\Rightarrow \begin{cases} \frac{1}{n}\sum_{i=1}^{n} x_i^2 a + x_i b = \frac{1}{n}\sum x_i y_i \\ \frac{1}{n}\sum_{i=1}^{n} x_i a + b = \frac{1}{n}\sum_{i=1}^{n} y_i \end{cases}$$

$$\Rightarrow \begin{cases} \left(\sum_{i=1}^{n} x_i^2\right)a + \left(\sum_{i=1}^{n} x_i\right)b = \sum_{i=1}^{n} x_i y_i \quad \text{--- (1)} \\ \left(\sum_{i=1}^{n} x_i\right)a + nb = \sum_{i=1}^{n} y_i \quad \text{--- (2)} \end{cases}$$

(1) & (2) linear system of Eqns in a & b. & solve for a & b. ✓

Using second derivative test on the above function, we can show (though it is hard) it is a minimum.

least square error: more general.

Example: (1) For best exponential fit

$$y = ce^{ax} \Rightarrow \underbrace{\ln y = ax + \ln c}_{\text{linear fit.}}$$

(2) Quadratic fit: $y = ax^2 + bx + c$

$$\underbrace{J(a,b,c)}_{\text{minimize}} = \sum_{i=1}^{n} \left(y_i - (ax_i^2 + bx_i + c)\right)^2$$

Idea to be understood:

* All these best fit problems are minimization problems. (Finding minimum of multivariable functions!)

From the critical point to conclude: local max, min or saddle point.

Second derivative test.

Example: Consider a quadratic function in 2 variables

$$f = ax^2 + bxy + cy^2$$

With some parameter value taken

$$f = x^2 + 2xy + 3y^2. \rightarrow \text{Critical point } f_x = 0 \ \& \ f_y = 0 \Rightarrow (0,0)$$

$$= \underline{(x+y)^2 + y^2} \geq 0$$

From this function; $(0,0)$ is local minimum.

In general (For quadratic $a \neq 0$)

$$f = a\left(x^2 + \frac{b}{a}xy\right) + cy^2$$

$$= a\left\{\left(x + \frac{by}{2a}\right)^2 - \frac{b^2}{4a^2}y^2\right\} + cy^2$$

$$= a\left(x+\frac{b}{2a}y\right)^2 + \left(c-\frac{b^2}{4a}\right)y^2$$

$$f = \frac{1}{4a}\Big[\underbrace{\underbrace{4a^2\left(x+\frac{b}{2a}y\right)^2}_{\text{Term 1}} + \underbrace{(4ac-b^2)y^2}_{\text{term 2}}}_{\text{sum/ difference of two squares.}}\Big]$$

We saw example

$f = x^2+y^2 \rightarrow$ local min. at $(0,0)$ (sum of 2 squares)

$f = y^2-x^2 \rightarrow$ Saddle point at $(0,0)$ (Difference of 2 squares).

3 cases: (1) $4ac-b^2<0$

term 1 ≥ 0 ; term 2 <0 } $\Rightarrow$ Saddle point.

"Similar to difference of two squares."

Ref: Vector Calculus by Michael Corral.

$f = x^2 + y^2$

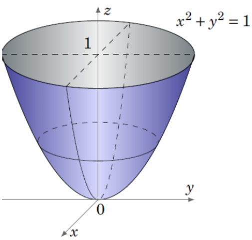


Figure 3.5.2 $z = x^2 + y^2$

function with Global min at (0,0)

$f = x^2 - y^2$

function with saddle point at (0,0)

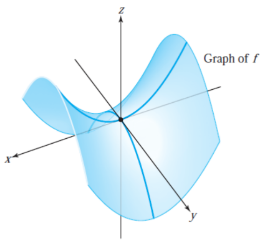


figure **3.3.3** A function of two variables with a saddle point.

Ref: Vector Calculus (6th Ed) by Jerrold E Marsden Anthony Tromba.

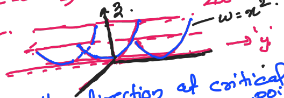
భారతీయ సాంకేతిక విజ్ఞాన సంస్థ హైదరాబాద్
भारतीय प्रौद्योगिकी संस्थान हैदराबाद
Indian Institute of Technology Hyderabad

Case 2

$4ac - b^2 = 0.$: degenerate case : where function varies in a certain direction. (here depends on 'a')

↳ Similar to quadratic determinant; when $b^2 - 4ac = 0$ ⇒ Soln ~ $\frac{-b}{2a}$

Ex: $w = x^2$.



"Function does not change its value except its direction at critical points"

Case 3: $4ac - b^2 > 0.$

⇒ term ① ≥ 0; term ② ≥ 0 ⇒ Sum of two squares.

$$\Rightarrow f = \frac{1}{4a}\underbrace{\left[(\;)^2 + (\;)^2\right]}_{\geq 0}.$$

∴ $f(0,0)$ is either maxima or minima depends on value of 'a'

If $a > 0$ ⇒ minimum. (∵ f is always +ve & greater than value at (0,0))

If $a < 0$ ⇒ maximum. (f is always -ve & less than '0').

Now let us take general function: & its second derivative.

$\frac{\partial^2 f}{\partial x^2}$; $\frac{\partial^2 f}{\partial y^2}$; $\frac{\partial^2 f}{\partial x \partial y}$; $\frac{\partial^2 f}{\partial y \partial x}$; $\frac{\partial^2 f}{\partial x \partial y} = \frac{\partial^2 f}{\partial y \partial x}$ if they exist.

$\parallel$ f_{xx} $\quad$ $\parallel$ f_{yy} $\quad$ f_{xy} $\quad$ f_{yx}.

Second derivative test!

At critical point (x_0, y_0) of function $f(x,y)$

Let $A = f_{xx}(x_0, y_0)$; $B = f_{xy}(x_0, y_0)$; $C = f_{yy}(x_0, y_0)$

Case 1: If $AC - B^2 > 0$; (i) $A > 0 \rightarrow$ local minimum,

(ii) $A < 0 \rightarrow$ local maximum.

Case 2: $AC - B^2 < 0 \rightarrow$ Saddle point.

Case 3: $AC - B^2 = 0 \rightarrow$ Cannot conclude. (degenerate case)

Example:

Find all local maxima & minima of $f(x,y) = xy - x^3 - y^2$.

Soln: Critical points:

$\frac{\partial f}{\partial x} = 0 \Rightarrow y - 3x^2 = 0 \Rightarrow y = 3x^2 = 0$

$\frac{\partial f}{\partial y} = 0 \Rightarrow x - 2y = 0 \Rightarrow x - 6x^2 = 0 \Rightarrow x = 0;\ 1/6$.

$6x^2 = x$

$x = 1/6$.

$y = 0;\ y = 3 \times \frac{1}{36} = \frac{1}{12}$.

Critical points: $(0,0)$, $(1/6, 1/12)$

$(0,0) \Rightarrow$

$\frac{\partial^2 f}{\partial x^2}(0,0) = -6x|_{00} = 0 = A$

$\frac{\partial^2 f}{\partial x \partial y}(0,0) = 1 = B$.

$\frac{\partial^2 f}{\partial y^2} = -2 = C$.

$AC - B^2 = -1 < 0$.

$\Rightarrow$ saddle point.

$\left(\frac{1}{6}, \frac{1}{12}\right) \Rightarrow$

$A = -6 \times \frac{1}{6} = -1$.

$B = 1$.

$C = -2$.

$AC - B^2 = 2 - 1 = 1 > 0$.

$A = -1 < 0 \Rightarrow$ local maximum.