# Multivariable Calculus Review

**Key Idea:** Approximate by simple functions such as linear or quadratic ones.

**Def:** $f(x,y)$ is differentiable at $\vec{x}_0 = \begin{bmatrix} x_0 \\ y_0 \end{bmatrix}$ if there ex. a linear function $L_{f,\vec{x}_0}(\vec{x})$ with

$$\lim_{\vec{x} \to \vec{x}_0} \frac{|f(\vec{x}) - L_{f,\vec{x}_0}(\vec{x})|}{|\vec{x} - \vec{x}_0|} = 0.$$

If such $L(\vec{x})$ exists, how do we find it?

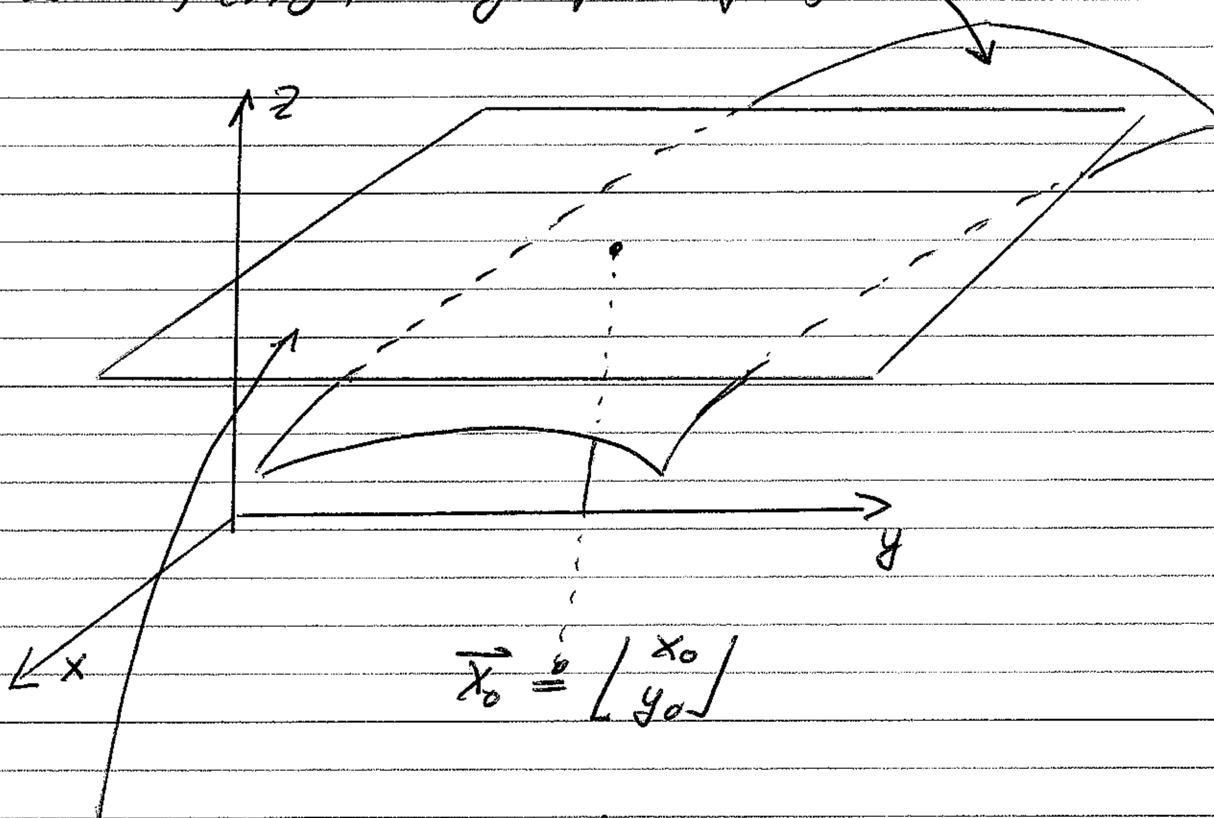
$$L_{f,\vec{x}_0}(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0).$$

$$\nabla f(\vec{x}_0) = \begin{bmatrix} \frac{\partial f}{\partial x}(\vec{x}_0) \\ \frac{\partial f}{\partial y}(\vec{x}_0) \end{bmatrix}.$$ gradient of $f$ at $\vec{x}_0$.

We say that $L_{f,\vec{x}_0}(\vec{x})$ is the best linear approximation of $f(\vec{x})$ at the point $\vec{x}_0$.

## Geometric Picture

$z = f(x,y)$ graph of $f$.




graph of linear function

$$z = L_{f,\vec{x}_0}(\vec{x}) = f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x} - \vec{x}_0)$$

This is the plane, tangent to the graph of $f$ at the point

$$\begin{bmatrix} x_0 \\ y_0 \\ f(x_0, y_0) \end{bmatrix}.$$

Directional derivative

(1) $\frac{d}{dt} f(\vec{x}_0 + t\vec{v}) = \nabla f(\vec{x}_0) \cdot \vec{v}$.

Interpretation: Consider the function

$g(t) = f(\vec{x}_0 + t\vec{v})$.

$g'(0)$ is the rate of change of $g(t)$ at $t=0$. Thus $\nabla f(\vec{x}_0) \cdot \vec{v}$ is the rate of change of the height of the graph of $z = f(\vec{x})$. as you move in the $\vec{x}$-plane starting from $\vec{x}_0$ in the direction $\vec{v}$.

If $\vec{x}(t)$ is a curve then

(2) $\frac{d}{dt} f(\vec{x}(t))\Big|_{t=t_0} = \nabla f(\vec{x}(t_0)) \cdot \vec{x}'(t)$.

From (1) we glean the fact that the gradient always points in the direction of largest increase of the function $f$.

Recall that a <u>level curve</u> is the set

$$\{\vec{x} \text{ in } \mathbb{R}^2 : f(\vec{x}) = \text{const.}\}.$$

If $\vec{x}(t)$ is the parametization of the level curve $f(\vec{x}) = f(\vec{x}_0)$ and $\vec{x}(0) = \vec{x}_0$, then from (2) we get.

$$0 = \frac{d}{dt} f(\vec{x}(t))\Big|_{t=0} = \nabla f(\vec{x}_0) \cdot \vec{x}\,'(0).$$

The gradient is always perpendicular to the level curves.

## Recall from 1-dim Calculus

$g(t)$ nice function, as many times differentiable as you like.

Then

$$g(t) = g(0) + g'(0)t + \frac{1}{2}g''(0)t^2 + \dots + \frac{1}{n!}g^{(n)}(0)t^n$$
$$+ \frac{1}{(n+1)!}g^{(n+1)}(c)\,t^{n+1}$$

where $c$ is some number between $0$ and $t$. Thus $c$ depends on $t$.

In particular we have that

$$g(t) - \left[g(0) + g'(0)t + \frac{1}{2}g''(0)t^2\right] = \frac{1}{3!}g'''(c)t^3$$

Hence

$$\lim_{t \to 0} \frac{\left|g(t) - \left[g(0) + g'(0)t + \frac{1}{2}g''(0)t^2\right]\right|}{t^2} = 0 \quad (3)$$

and we can say that $g(0) + g'(0)t + \frac{1}{2}g''(0)t^2$ is the best quadratic approximation of $g$ at $t$.!

Consider now a function in two variables, $f(x,y)$. Pick any vector $\overline{v}$.

Set $g(t) = f(\overline{x}_0 + t\overline{v})$

Then $g(0) = f(\overline{x}_0)$

$g'(0) = \nabla f(\overline{x}_0) \cdot \vec{v}$ and

$g''(0) = \vec{v} \cdot H_f(\overline{x}_0)\overline{v}$ where

$$H_f(\overline{x}_0) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2}(\overline{x}_0) & \frac{\partial^2 f}{\partial y \partial x}(\overline{x}_0) \\ \frac{\partial^2 f}{\partial x \partial y}(\overline{x}_0) & \frac{\partial^2 f}{\partial y^2}(\overline{x}_0) \end{bmatrix}$$

is the Hessian matrix.

For "nice" functions $f$ we always have

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

"Nice" means that the second partial derivatives of $f$ exist and are continuous.

Set $t\vec{v} = \vec{x} - \vec{x}_0$

Hence, the best quadratic approximation of $f$ at $\vec{x}_0$ is given by

$$Q_{f,\vec{x}_0}(\vec{x}) := f(\vec{x}_0) + \nabla f(\vec{x}_0) \cdot (\vec{x}-\vec{x}_0) + \frac{1}{2}(\vec{x}-\vec{x}_0) \cdot H_f(\vec{x}_0)(\vec{x}-\vec{x}_0)$$

i.e.,

$$\lim_{\vec{x} \to \vec{x}_0} \frac{|f(\vec{x}) - Q_{f,\vec{x}_0}(\vec{x})|}{|\vec{x}-\vec{x}_0|^2} = 0.$$

This follows from (3)!

The next question is how to <u>visualize</u> the quadratic approximation.

The linear approximation is easy to visualize, since the graph of this linear function is just a plane.

The trick is to choose a new coordinate system so that the function $(\vec{x}-\vec{x}_0) \cdot H_f(\vec{x}_0)(\vec{x}-\vec{x}_0)$ looks as simple

as possible. Let us not worry about $f(\vec{x}_0) + \nabla f(\vec{x}_0)\cdot(\vec{x}-\vec{x}_0)$ but only consider the quadratic part.
By a translation, we may always assume that $\vec{x}_0 = 0$. Hence we have to study

$$\underline{\underline{\vec{x} \cdot H_f(0)\, \vec{x}}}.$$

$H_f(0)$ is a symmetric matrix and hence it can be diagonalized, i.e., there exist two vectors $\vec{v}_1, \vec{v}_2$ (which we normalize such that $|\vec{v}_1| = |\vec{v}_2| = 1$) so that

$$H_f(0)\,\vec{v}_1 = \mu_1 \vec{v}_1$$

$$H_f(0)\,\vec{v}_2 = \mu_2 \vec{v}_2$$

$\vec{v}_1, \vec{v}_2$ are called <u>eigenvectors</u> and $\mu_1, \mu_2$ are called the eigenvalues.

It is important to note that always

$$\vec{v}_1 \cdot \vec{v}_2 = 0$$

that is, the vectors are perpendicular.

Now

$$H_f(0)\,[\vec{v}_1, \vec{v}_2] = [H_f(0)\vec{v}_1, H_f(0)\vec{v}_2]$$

$$\underset{\text{column vectors}}{\uparrow\ \uparrow} \qquad = [\mu_1\vec{v}_1, \mu_2\vec{v}_2]$$

$$= [\vec{v}_1, \vec{v}_2]\begin{bmatrix}\mu_1 & 0\\ 0 & \mu_2\end{bmatrix}.$$

Hence if we set $[\vec{v}_1, \vec{v}_2] = V$, then

$H_f(0) = V\begin{bmatrix}\mu_1 & 0\\ 0 & \mu_2\end{bmatrix}V^{-1}$ where

$V^{-1} = \begin{bmatrix}\vec{v}_1\\ \vec{v}_2\end{bmatrix}$. It is easy to check that

$$\begin{bmatrix}\vec{v}_1\\ \vec{v}_2\end{bmatrix}[\vec{v}_1, \vec{v}_2] = \begin{bmatrix}\vec{v}_1\cdot\vec{v}_1 & \vec{v}_1\cdot\vec{v}_2\\ \vec{v}_2\cdot\vec{v}_1 & \vec{v}_2\cdot\vec{v}_2\end{bmatrix}$$

$= \begin{bmatrix}1 & 0\\ 0 & 1\end{bmatrix}$ since $|\vec{v}_1| = |\vec{v}_2| = 1$ and $\vec{v}_1\cdot\vec{v}_2 = 0$.

Thus, since $\begin{bmatrix}\vec{v}_1\\ \vec{v}_2\end{bmatrix} = V^T$ we have that $V^{-1} = V^T$, i.e., $V$ is an orthogonal matrix

Hence our quadratic function can be written as

$$\vec{x} \cdot H_f(0)\vec{x} = \vec{x} \cdot V\begin{bmatrix}\mu_1 & 0 \\ 0 & \mu_2\end{bmatrix}V^T\vec{x}$$

$$= (V^T\vec{x}) \cdot \begin{bmatrix}\mu_1 & 0 \\ 0 & \mu_2\end{bmatrix}(V^T\vec{x})$$ since quite generally

$$\vec{x} \cdot A\vec{y} = A^T\vec{x} \cdot \vec{y}$$

$x \in \mathbb{R}^m$, $A$ $m \times n$ matrix and $y \in \mathbb{R}^n$.

If we set $\vec{y} = V^T\vec{x}$ we have

$$\vec{x} \cdot H_f(0)\vec{x} = \vec{y} \cdot \begin{bmatrix}\mu_1 & 0 \\ 0 & \mu_2\end{bmatrix}\vec{y}$$

$$= \underline{\underline{\mu_1 y_1^2 + \mu_2 y_2^2}}$$ if we set $\vec{y} = \begin{bmatrix}y_1 \\ y_2\end{bmatrix}$.

The level curves of this function are easy to ~~visualize~~ visualize. They are ellipses if $\mu_1 > 0$, $\mu_2 > 0$ or $\mu_1 < 0$ and $\mu_2 < 0$. They are hyperbolas if $\mu_1$ and $\mu_2$ have opposite sign.

How are the $\vec{y}$ picture and the $\vec{x}$ picture related?

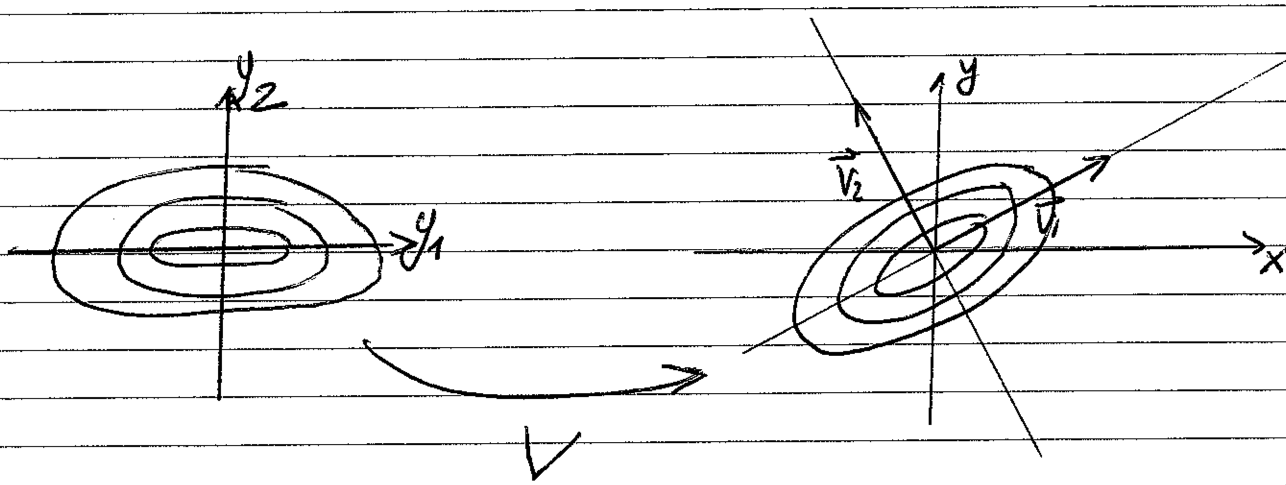
$\vec{y} = V^T\vec{x} = V^{-1}\vec{x}$ and hence

$\vec{x} = V\vec{y} = [\vec{v}_1, \vec{v}_2]\vec{y}$

If $\vec{y} = \begin{bmatrix}1\\0\end{bmatrix}$ then $\vec{x} = \vec{v}_1$ and if

$\vec{y} = \begin{bmatrix}0\\1\end{bmatrix}$ then $\vec{x} = \vec{v}_2$.

Thus the $y_1$ axis is mapped to the $\vec{v}_1$ direction and the $y_2$ axis is mapped to the $\vec{v}_2$ direction.




You just rotate the $\vec{y}$-picture by the rotation $V$ ($V$ can always be chosen to be a rotation.

Example:

Suppose $H_f(0) = \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix}$.

Calculate first eigenvalues:

$$\det \begin{bmatrix} 3-\mu & 1 \\ 1 & 3-\mu \end{bmatrix} = \mu^2 - 6\mu + 8 = 0$$

$\mu_1 = 4 \quad \mu_2 = 2$

Now calculate the eigenvector:

$$\begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\vec{v}_1} = 4 \underbrace{\begin{bmatrix} a \\ b \end{bmatrix}}_{\vec{v}_1} \text{ or}$$

$3a + b = 4a$ or $b = a$

$a + 3b = 4b$ or $a = b$

$$\vec{v}_1 = \underbrace{\frac{1}{\sqrt{2}}}_{\text{normalization}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

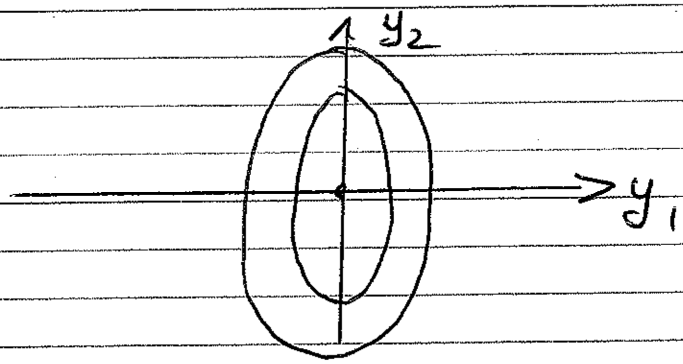
normalization.

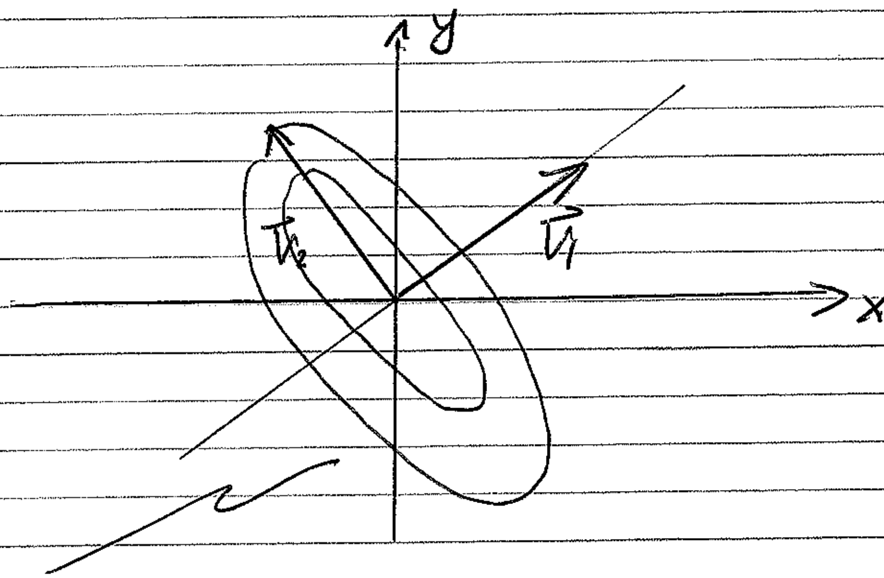
$\vec{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$ because $\vec{v}_1 \cdot \vec{v}_2 = 0$!

Hence we can draw pictures:

The level curves of

$4y_1^2 + 2y_2^2$ are ellipses.




these are the level sets of the function

$\vec{x} \cdot \begin{bmatrix} 3 & 1 \\ 1 & 3 \end{bmatrix} \vec{x}$.

## Critical points

**Def**: $\vec{x}_0$ is a critical point for $f$ if

$$\nabla f(\vec{x}_0) = 0 .$$

At local maxima $\vec{x}_M$ and local minima $x_m$

we have $\nabla f(\vec{x}_m) = \nabla f(\vec{X}_M) = 0$ .

If we want to decide the type of critical point $\vec{x}_0$ we look at the quadratic approximation of $f(\vec{x})$ near this point $\vec{x}_0$ .

$$(\vec{x}-\vec{x}_0) \cdot H_f(\vec{x}_0)(\vec{x}-\vec{x}_0) .$$

Now we perform the diagonalisation as before. We can say right away:

If $\mu_1$ and $\mu_2 > 0$ , local minimum

If $\mu_1$ and $\mu_2 < 0$ , local maximum

If $\mu_1$ and $\mu_2$ have opposite sign $\vec{x}_0$ is a saddle point.

Example:

$f(x,y) = x^3 + y^3 + 3xy$

for crit. points.

$\nabla f = 3\begin{bmatrix} x^2+y \\ y^2+x \end{bmatrix}$ $\quad x^2+y=0 \;,\; y^2+x$

$x^2 = -y \quad x^4 + x = 0 \quad (x^3+1)x = 0$

$x=0 \Rightarrow y=0 . \quad x=-1 \Rightarrow y=-1$

Two critical points

$\vec{x}_0 = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$ and $\vec{x}_1 = \begin{bmatrix} -1 \\ -1 \end{bmatrix}$.

$H_f(\vec{x}) = \begin{bmatrix} 6x & 3 \\ 3 & 6y \end{bmatrix} = 3\begin{bmatrix} 2x & 1 \\ 1 & 2y \end{bmatrix}$

Hessian

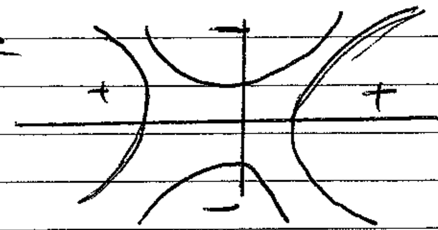
$H_f(\vec{x}_0) = 3\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \quad H_f(\vec{x}_1) = \begin{bmatrix} -6 & 3 \\ 3 & -6 \end{bmatrix}$

$= 3\begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$

$3\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$: eigenvalues: $3 \;, -3$

with eigenvectors $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix} \;, \frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

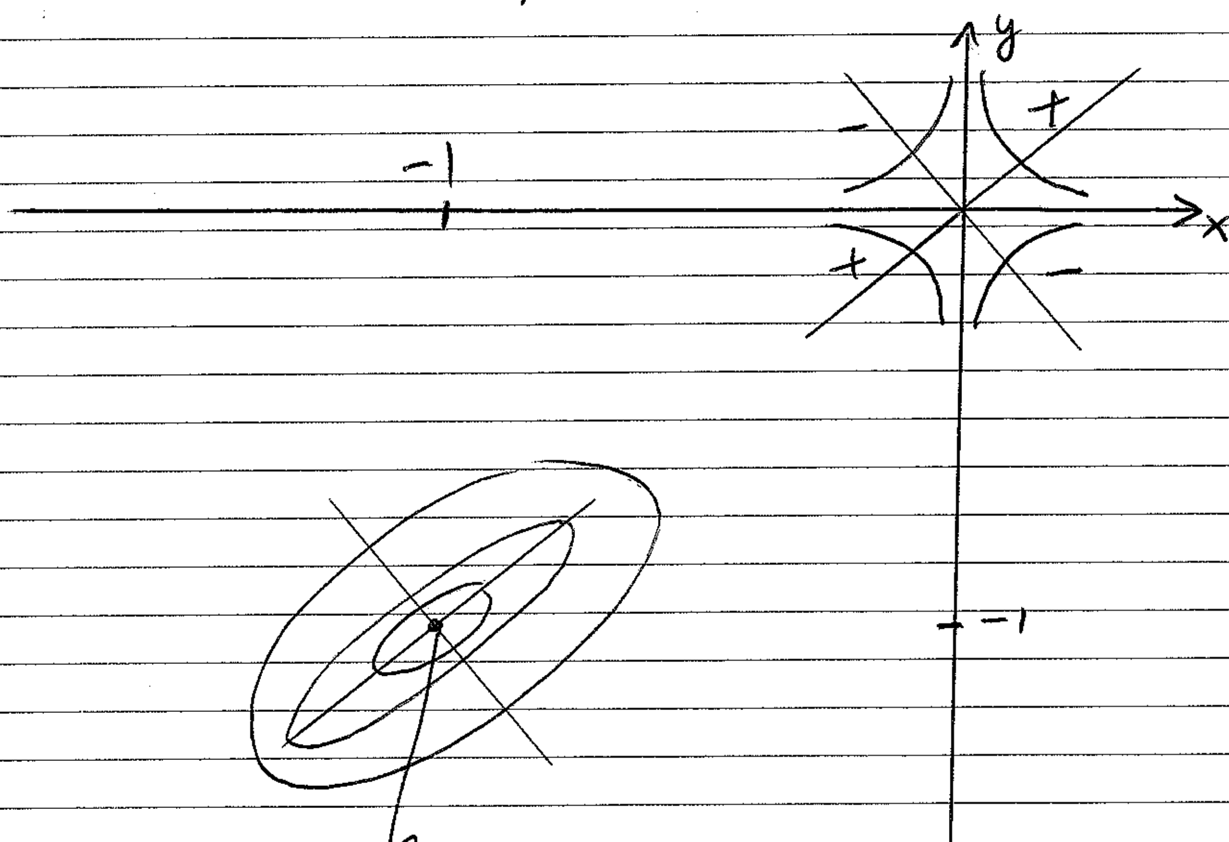
Thus the critical point is a saddle point. $\bar{y}$-plane picture




$3\begin{bmatrix} -2 & 1 \\ 1 & -2 \end{bmatrix}$: eigenvalues: $-3$, $-9$ local min.

with eigenvectors: $\frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$, $\frac{1}{\sqrt{2}}\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Hence we have the following picture near the critical points.




$f(-1,-1) = 1$

local min