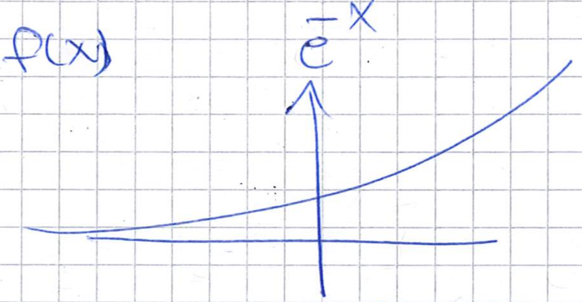
(F7)

(1) $\|r(x)\|_2 \geq 0$. Special structure $r(x)^T r(x)$

---

(2) <u>proof</u> Pick $x_0 \in \mathbb{R}^n$ and set $M = f(x_0) + 1$.
Since $f(x)$ is coercive there is an α so $f(x) \geq M$ if $\|x\| > \alpha$. So $f(x)$ is cont. on a closed set $D = \{x : \|x\|_2 \leq \alpha\} \Rightarrow$ has a minimum on D. This is a global minima.



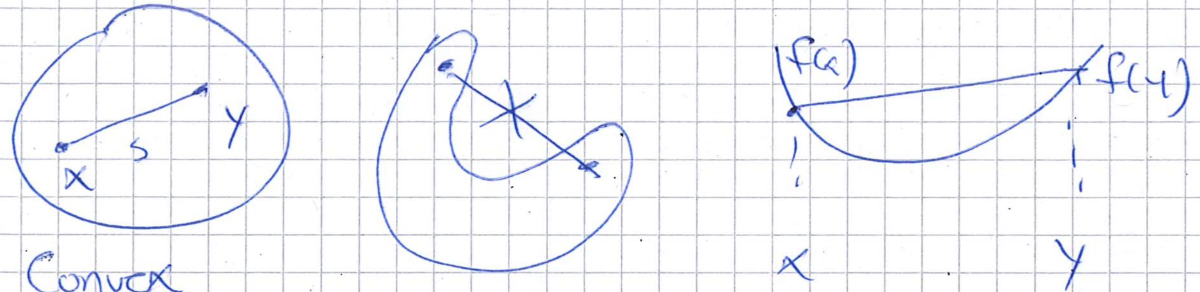
not coercive

$x_1^2 - 2x_1x_2 + 3x_2^2$
$= (x_1 - x_2)^2 + 2x_2^2$
is coercive

$x_1^2 + x_1 e^{-x_2}$
not coercive since $x_1 = 0 \Rightarrow f = 0$

---

(3)



Convex

<u>proof</u> If $f(x)$ has two different global minima $x \neq y$ and $f(x) = f(y)$ then

$$f\left(\frac{x+y}{2}\right) < \frac{1}{2}f(x) + \frac{1}{2}f(y) = f(x)$$

so neither x or y were global minima.

④ Taylor series $f: \mathbb{R}^2 \to \mathbb{R}$

$$f(x,y) = f(x_0,y_0) + f'_x (x-x_0) + f'_y (y-y_0) +$$

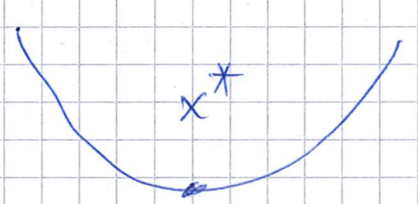
$$\frac{1}{2} f''_{xx} (x-x_0)^2 + \frac{2}{2} f''_{xy} (x-x_0)(y-y_0) + \frac{1}{2} f''_{yy} (y-y_0)^2 + \dots$$

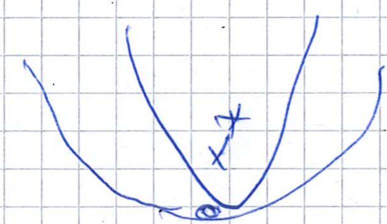
$$= f(x_0,y_0) + \left(f'_x, f'_y\right) \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix} + \frac{1}{2} \left(x-x_0 \;\; y-y_0\right) \begin{pmatrix} f''_{xx} & f''_{xy} \\ f''_{xy} & f''_{yy} \end{pmatrix} \begin{pmatrix} x-x_0 \\ y-y_0 \end{pmatrix}$$

$$= \left\{ x = \begin{pmatrix} x \\ y \end{pmatrix} \right\} = f(x_0) + \nabla f(x_0)^T (x-x_0) + \frac{1}{2} (x-x_0)^T H_f (x-x_0)$$

$$+ \dots$$

The matrix H_f is called <u>Hessian</u>





$f'(x^*) = 0$ and

$f''(x^*) \geq 0$

$\Rightarrow$ local minima

$\nabla f(x^*) = 0$ and

$H_f(x^*)$ pos. def

$\Rightarrow$ local minima.

(5) $$f(x) \approx f(x^{(k)}) + \left(\nabla f(x^{(k)})\right)^T (x - x^{(k)}) +$$
$$+ \frac{1}{2}(x - x^{(k)})^T H_f (x - x^{(k)})$$

$$\Rightarrow \nabla f(x) \approx \nabla f(x^{(k)}) + H_f \cdot (x - x^{(k)}) = 0$$

$$\Rightarrow H_f \cdot s^{(k)} = -\nabla f(x^{(k)})$$

This is Newton's method.

(6) $$f(x) = \frac{1}{2} r^T r = \frac{1}{2} \sum_{i=1}^{m} r_i^2(x)$$

$$\Rightarrow \nabla f = \frac{1}{2} \sum_{i=1}^{m} \nabla (r_i^2(x)) = \sum \nabla r_i(x) \cdot r_i(x)$$

$$= \begin{pmatrix} \nabla r_1 \\ \vdots \\ \nabla r_n \end{pmatrix} r(x) = J_r^T \cdot r(x)$$

The product rule gives

$$(f)'' = \begin{pmatrix} \nabla r_1 \\ \vdots \\ \nabla r_m \end{pmatrix}' r(x) + \begin{pmatrix} \nabla r_1 \\ \vdots \\ \nabla r_m \end{pmatrix} \underbrace{(r(x))'}_{J_r}$$

$$= \sum_{i=1}^{m} H_{r_i} \cdot r_i + J_r^T J_r$$

(7) Newtons method is

$$H_f s^{(k)} = -\nabla f(x^{(k)}) = -J_r^T r(x^{(k)})$$

Approximate $H_f \approx J_r^T J_r$ and obtain

$$J_r^T J_r s^{(k)} = -J_r^T r(x^{(k)})$$

which is normal equations for

$$\min_{s^{(k)}} \| J_r s^{(k)} + r(x^{(k)}) \|_2$$

---

(8) We get $r(x) = \begin{pmatrix} x_1 - 2 \\ x_1 e^{x_2} - 0.7 \\ x_1 e^{2x_2} - 0.3 \\ x_1 e^{3x_2} - 0.1 \end{pmatrix} : \mathbb{R}^2 \to \mathbb{R}^4$

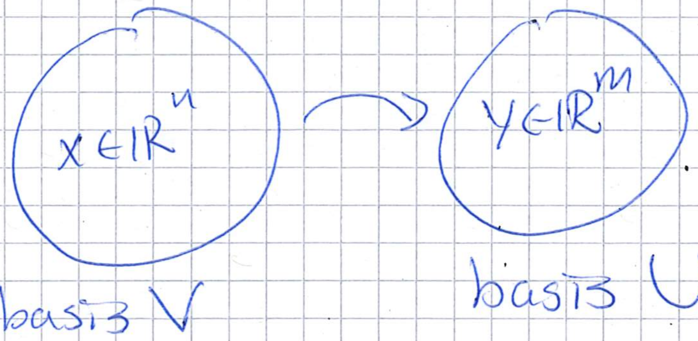
and

$$J_r(x) = \begin{pmatrix} 1 & 0 \\ e^{x_2} & x_1 e^{x_2} \\ e^{2x_2} & 2x_1 e^{2x_2} \\ e^{3x_2} & 3x_1 e^{3x_2} \end{pmatrix} \in \mathbb{R}^{4\times 2}$$

$$\begin{array}{cc} \uparrow & \uparrow \\ \frac{\partial r}{\partial x_1} & \frac{\partial r}{\partial x_2} \end{array}$$

---

(9) $A: \mathbb{R}^n \to \mathbb{R}^m$



If

$$x = Ix = VV^T x = Vq = (v_1, v_2, \ldots, v_n) q = \sum_{i=1}^{n} q_i v_i$$

then $(q_1, \ldots, q_n)$ are the coordinates of x in the basis V //

---

(10) <u>proof</u> $A = U \Sigma V^T \Rightarrow AV = U\Sigma$

or
$$A(v_1, \ldots, v_n) = (u_1, \ldots, u_m) \begin{pmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_n \\ & & 0 & \end{pmatrix}$$

$$\Leftrightarrow (Av_1, \ldots, Av_n) = (\sigma_1 u_1, \ldots, \sigma_n u_n)$$

So
$$Av_i = \sigma_i u_i$$ //

Similar $A^T = V \Sigma^T U^T$ gives $A^T u_i = \sigma_i v_i$ //

<u>Proof</u> $(m > n)$

$$A = U \Sigma V^T = (u_1, \ldots, u_m) \begin{pmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_n \\ & 0 & \end{pmatrix} \begin{pmatrix} v_1^T \\ \vdots \\ v_n^T \end{pmatrix} =$$

$$(u_1, \ldots, u_m) \begin{pmatrix} \sigma_1 v_1^T \\ \vdots \\ \sigma_n v_n^T \end{pmatrix} = \sum_{i=1}^{n} u_i \sigma_i v_i^T$$ //

(11) <u>proof</u> We have $A = \sum_{i=1}^{n} \sigma_i u_i v_i^T$.

Write x using the basis V gives

$$x = \sum_{i=1}^{n} (v_i^T x) v_i \Rightarrow Ax = \sum_{i=1}^{n} (v_i^T x) A v_i = \sum_{i=1}^{n} \sigma_i (v_i^T x) u_i$$

and write b in the basis U

$$b = \sum_{i=1}^{n} (u_i^T b) u_i$$

So we get

$$Ax = b \Leftrightarrow \sum_{i=1}^{n} \sigma_i (v_i^T x) u_i = \sum_{i=1}^{n} (u_i^T b) u_i$$

$$\Rightarrow (v_i^T x) = \frac{u_i^T b}{\sigma_i}$$

So the solution is

$$x = \sum_{i=1}^{n} \frac{u_i^T b}{\sigma_i} v_i$$

(12) <u>proof</u>

$$\|A\|_2^2 = \|U\Sigma V^T\|_2^2 = \begin{Bmatrix} U, V \\ \text{orthogonal} \end{Bmatrix}$$

$$= \|\Sigma\|_2^2 = \max_{\|y\|_2=1} \|\Sigma y\|_2^2 = \max_{\|y\|_2=1} \sum_{i=1}^{n} \sigma_i^2 y_i^2$$

$$\leq \sigma_1^2 \max_{\|y\|_2=1} \sum_{i=1}^{n} y_i^2 = \sigma_1^2$$

with equality for $y = e_1$ //