Théorème des extrema liés - Lagrange multipliers

Primal Pb : $x^* = \min_{x \in \mathcal{D}_a} f(x)$, $\mathcal{D}_a = \{x, C(x) = 0\}$

Let f and C in C^1, and $x^\star$ a local minimizer of f satisfying

$f : \mathbb{R}^n \to \mathbb{R}$

$C : \mathbb{R}^n \to \mathbb{R}^m$

$$C(x^\star) = 0 \quad \text{primal feasability}$$

rank $\nabla J C(x^*) = m$

If the constraints are qualified, there exists a vector of Lagrange multipliers $y^\star \in \mathbb{R}^m$ s. t.

$\nabla_x \ell(x^*, y^*) = 0_{\mathbb{R}^n}$

$$\nabla f(x^\star) + \sum_{i=1}^{m} y_i^\star \nabla C_i(x^\star) = 0 \quad \text{dual feasability}$$

look for (x^*, y^*) such that $\begin{cases} C(x^*) = 0_{\mathbb{R}^m} \\ \nabla_x \ell(x^*, y^*) = 0_{\mathbb{R}^n} \end{cases}$

$x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$

- Linear constraints special case
- $n = 2$, $m = 1$ special case

Linear constraints special case

$\inf_{c(x)=0} f(x)$

$$c(x) = (\langle c_1, x\rangle, \ldots, \langle c_m, x\rangle)^\perp$$ qualified hypothesis

$(c_i)_{i=1,\ldots,m}$ independant vectors family in R^n

$$\mathcal{D}_a = K = \{x, \langle c_i, x\rangle = 0, i = 1, \ldots, m\} = \text{admissible set}$$

$$\inf_{x\in K} f(x) \Leftrightarrow \inf_{\alpha\in\mathbb{R}^p} g(\alpha) = g(\alpha^*) \Rightarrow \nabla g(\alpha^*) = 0$$

with $g(\alpha) = f\left(\sum_{i=1}^{p} \alpha_i k_i\right)$, $(k_i)_{i=1,\ldots,p}$, basis of K

$\hookrightarrow n-m$

$$g(\alpha^*) = f\left(\sum_{i=1}^{p} \alpha_i k_i\right) \qquad \nabla g(\alpha^*) = \left(\frac{\partial g}{\partial \alpha_i}\right)_{i=1,\dots p}$$

$$\frac{\partial g}{\partial \alpha_i}(\alpha^*) = \langle k_i, \nabla f(\textstyle\sum \alpha_i^* k_i) \rangle = 0 \quad \forall i \dots$$

$$\nabla f(x^*) \in \overset{\text{v.s}}{ev}\{(k_i)\}^{\perp} = \overset{\text{v.s}}{ev}\{(c_i)^{\perp}\}^{\perp} = \overset{\text{v.s}}{er}\{(c_i)\} \Rightarrow$$

$$\exists \lambda_i, \; \nabla f(x^*) = \sum_{i=1}^{m} \lambda_i c_i$$

$$\nabla f(x^*) + \langle y_i^*, c(x^*) \rangle = 0 \quad \text{choose } y_i^* = -\lambda_i$$

v.s vector space generated by a family of vectors.

e.v espace vectoriel

Special case $n = 2$, $m = 1$

$C : \mathbb{R} \to \mathbb{R}$

$\inf\limits_{C(x) = 0} f(x)$

Qualification condition for one single constraint $m = 1$:
$\nabla_x c_1(x^\star) \neq 0$, we can suppose $\partial_{x_2} c_1(x^\star) \neq 0$.
Implicit function theorem : $\exists V_1 \times V_2$ containing $x^\star$ and φ unique and differentiable in $x^\star$ s. t. $\forall x_1 \in V_1$ $c_1([x_1, \varphi(x_1)]) = 0$ and $x_2^\star = \varphi(x_1^\star)$ with

$$\varphi'(x_1) = \frac{-1}{\partial_{x_2} c_1(x)} \partial_{x_1} c_1(x).$$

Proof

$$\inf_{c_1(x)=0} f(x) \quad \Leftrightarrow \quad \inf_{x_1 \in V_1} \tilde{f}(x_1), \qquad \text{with } \tilde{f}(x_1) = f([x_1, \varphi(x_1)])$$

First order ptimality conditions for $\tilde{f}$ (without constraints since V_1 is an open set)

$$\tilde{f}'(x_1^\star) = 0 \Leftrightarrow \frac{\partial f}{\partial x_1}([x_1^\star, \varphi(x_1^\star)]) + \varphi'(x_1^\star)\frac{\partial f}{\partial x_2}([x_1^\star, \varphi(x_1^\star)]) = 0.$$

$$y = -\frac{\partial_{x_2} f(x^\star)}{\partial_{x_2} c_1(x^\star)}$$

Exemple 1

$f : \mathbb{R}^2 \to \mathbb{R}$
$c : \mathbb{R}^2 \to \mathbb{R}$ $\quad c(x) = \|x\|^2 - 1$

D_a is the circle of center 0 and radius 1.

$$\inf_{x_1^2+x_2^2=1} x_1^4 + x_2^4 = f(x)$$

Resolution by changing variables in polar coordinates

Set $x_1 = \cos(\theta)$, $x_2 = \sin(\theta)$, problem (4) becomes $\inf_{\theta\in[0,2\pi]}((\cos\theta)^4 + (\sin\theta)^4)$ whose solution is obtained by finding the zero of the derivative:

$$4\cos\theta\sin\theta(-(\cos\theta)^2 + (\sin\theta)^2) = -2\sin(2\theta)\cos(2\theta) = 0,$$

4 local minima $(\pm\sqrt{2}/2, \pm\sqrt{2}/2)$, où $f(x) = 1/2$,
4 local maxima $\{(1,0),(0,1),(-1,0),(0,-1))\}$, où $f(x) = 1$.

Resolution using Lagrange multipliers

Suppose $f(x^*) = \min f(x)$ $\exists y^* \in \mathbb{R}$ s. t.

$\nabla_x \ell(x^*, y^*) = 0_{\mathbb{R}^2}$

$C(x) = x_1^2 + x_2^2 - 1$

$JC(x) = 2(x_1 \quad x_2)$

We seek $x^\star \in \mathbb{R}^2$ and $y^\star \in \mathbb{R}$ s. t. Rank $JC(x^*) = 1$

$C(x^*) = 0$

$$
\begin{aligned}
(x_1^\star)^2 + (x_2^\star)^2 &= 1 \\
4(x_1^\star)^3 + y^\star 2x_1^\star &= 0 \\
4(x_2^\star)^3 + y^\star 2x_2^\star &= 0
\end{aligned}
$$

$\ell(x, y) = f(x) + \langle y, C(x) \rangle = x_1^4 + x_2^4 + y(x_1^2 + x_2^2 - 1)$

$\nabla_x \ell(x, y) = 4 \begin{pmatrix} x_1^3 \\ x_2^3 \end{pmatrix} + 2y \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}$

Resolution using Lagrange multipliers

We seek $x^\star \in \mathbb{R}^2$ and $y^\star \in \mathbb{R}$ s. t.

$$
\begin{aligned}
(x_1^\star)^2 + (x_2^\star)^2 &= 1 \\
2x_1(2x_1^2 + y) = 4(x_1^\star)^3 + y^\star 2x_1^\star &= 0 \\
2x_2(2x_2^2 + y) = 4(x_2^\star)^3 + y^\star 2x_2^\star &= 0
\end{aligned}
$$

	$x_1 = 0$	$y = -2x_1^2$
$x_2 = 0$	$\cancel{0 \notin \mathcal{D}_a}$	$x_1 = \pm 1$, $y = -2$, $f(x) = 1$
$y = -2x_2^2$	$x_2 = \pm 1$, $y = -2$, $f(x) = 1$	$y = -1$, $x_1 = \pm \sqrt{2}/2$, $x_2 = \pm \sqrt{2}/2$, $f(x) = \frac{1}{2}$

Resolution using Lagrange multipliers

We seek $x^\star \in \mathbb{R}^2$ and $y^\star \in \mathbb{R}$ s. t.

$$\begin{aligned} (x_1^\star)^2 + (x_2^\star)^2 &= 1 \\ 4(x_1^\star)^3 + y^\star 2x_1^\star &= 0 \\ 4(x_2^\star)^3 + y^\star 2x_2^\star &= 0 \end{aligned}$$

	$x_1^\star = 0$	$y^\star = -2(x_1^\star)^2$
$x_2^\star = 0$	$(x_1^\star)^2 + (x_2^\star)^2 \neq 1$	$(x_1^\star)^2 = 1$ et $y^\star = -2$ $f(x^\star) = 1$
$y^\star = -2(x_2^\star)^2$	$(x_2^\star)^2 = 1$ et $y^\star = -2$ $f(x^\star) = 1$	$(x_1^\star)^2 = (x_2^\star)^2 = 1/2$ et $y^\star = -1$, $f(x^\star) = 1/2$

Exemple 2

1) check that C qualified $\operatorname{Rank} JC(x) = 1$
2) check that x^* exists
3) apply the LM theorem

$\mathcal{D}_a = \{ x,\ x_1 + x_2 = 1 \} =$ line

Find

$f(x) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + c$

$$\inf_{C(x)=0} f(x) \quad f(x) = 3x_1^2 + 5x_2^2 - 3x_1x_2, \quad C(x) = x_1 + x_2 - 1$$

$= \left(x_1 - \frac{3}{2}x_2\right)^2 + 2x_1^2 + \left(5 - \frac{9}{4}\right)x_2^2 \geq 0 \longrightarrow +\infty$

$\|x\| \to +\infty$

- Lagrangian $\ell(x, y) = 3x_1^2 + 5x_2^2 - 3x_1x_2 + y(x_1 + x_2 - 1)$

f is a polynomial of degree 2 : therefor the Taylor expansion of degree 2 is exact

$f(x) = f(0) + \langle \nabla f(0), x\rangle + \frac{1}{2}\langle Hf(0)x, x\rangle$

$0 < \lambda_{min}, \lambda_{max}$

$\nabla f(x) = \begin{pmatrix} 6x_1 - 3x_2 \\ -3x_1 + 10x_2 \end{pmatrix}$

$Hf(x) = \begin{pmatrix} 6 & -3 \\ -3 & 10 \end{pmatrix}$ $\det Hf = 51 > 0$ and Trace $Hf = 16 > 0 \Rightarrow \lambda > 0$

$\lambda_{min}\|x\|^2 \leq \langle Ax, x\rangle \leq \lambda_{max}\|x\|^2$

Dim 1 $f(x) = f(0) + f'(0)x + \frac{1}{2} f''(0)x^2 + \frac{1}{6} f^{(3)}(0)x^3 + o(x^3)$

Exemple 2

Find

$$\inf_{C(x)=0} f(x) \quad f(x) = 3x_1^2 + 5x_2^2 - 3x_1x_2, \quad C(x^*) = x_1 + x_2 - 1 = 0$$

- Lagrangian $\ell(x, y) = 3x_1^2 + 5x_2^2 - 3x_1x_2 + y(x_1 + x_2 - 1)$
- Gradient $\nabla\ell(x^*, y^*) = \begin{pmatrix} 6x_1^* - 3x_2^* + y^* \\ -3x_1^* + 10x_2^* + y^* \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}$

Exemple 2

Find

$$\inf_{C(x)=0} f(x) \quad f(x) = 3x_1^2 + 5x_2^2 - 3x_1 x_2, \quad C(x) = x_1 + x_2 - 1$$

- Lagrangian $\ell(x, y) = 3x_1^2 + 5x_2^2 - 3x_1 x_2 + y(x_1 + x_2 - 1)$
- Gradient $\nabla\ell(x, y) = \begin{pmatrix} 6x_1 - 3x_2 + y \\ -3x_1 + 10x_2 + y \end{pmatrix}$
- If $f(x^\star) = \inf_{C(x)=0} f(x)$ then $\exists y^\star \in \mathbb{R}$ s.t. $\nabla\ell(x^\star, y^\star) = 0$

Exemple 2

Find

$$\inf_{C(x)=0} f(x) \quad f(x) = 3x_1^2 + 5x_2^2 - 3x_1x_2, \quad C(x) = x_1 + x_2 - 1$$

- Lagrangian $\ell(x, y) = 3x_1^2 + 5x_2^2 - 3x_1x_2 + y(x_1 + x_2 - 1)$
- Gradient $\nabla\ell(x, y) = \begin{pmatrix} 6x_1 - 3x_2 + y \\ -3x_1 + 10x_2 + y \end{pmatrix}$
- If $f(x^\star) = \inf_{C(x)=0} f(x)$ then $\exists y^\star \in \mathbb{R}$ s.t. $\nabla\ell(x^\star, y^\star) = 0$
- Plus the primal condition $C(x^\star) = 0$

Exemple 2

$$\begin{cases} 6x_1 - 3x_2 + y = 0 \\ -3x_1 + 10x_2 + y = 0 \\ x_1 + x_2 = 1 \end{cases}$$

to be solved as exercise.

- Solve the system of 3 equations to find $x^\star, y^*$

Exemple 2

Alternative method to minimize

$f(x) = 3x_1^2 + 5x_2^2 - 3x_1 x_2$

$\mathcal{D}_a : \{ x_1 + x_2 = 1 \}$

express $x_2 = 1 - x_1$ $g(x_1)$

linear

- Solve the system of 3 equations to find $x^\star$, y
- Other method ?

Second order optimality conditions

$\nabla_x \ell(x^*, y^*) = 0$

$H = H_x \ell(x^*, y^*)$ $\qquad A(x) = Jc(x)$

Let f and c in C^2, and $x^\star$ be a local minimizer of f verifying the constraints of equality $c(x^\star) = 0$. If the constraints are qualified, there exists a vector of Lagrange multipliers $y^\star \in \mathbb{R}^m$ such that

$$\langle s, H(x^\star, y^\star)s \rangle \geq 0 \quad \text{~~pour tout~~ for all } s \in \mathcal{N}$$

~~où~~ where

$$\mathcal{N} = \{s \in \mathbb{R}^n, A(x^\star)s = 0\}.$$

linear approximation of D_a

$D_a = \{x, c(x) = 0\}$

$c(x) = c(x^*) + Jc(x^*)(x - x^*) + o(\|x - x^*\|)$

$= Jc(x^*)(x - x^*) + o(\|x - x^*\|)$

$c(x) \simeq A(x^*)(x - x^*)$

Interpretation of Lagrange multipliers

The Lagrange multiplier y_i measures the sensitivity of the minimum $x^\star$ with respect to the corresponding constraint.

Initial primal and dual problems

$p^\star = \inf_{c(x)=0} f(x)$	$\sup_y g(y)$ avec $g(y) = \inf_x f(x) + y^t c(x)$

Perturbated primal and dual problems

$p^\star(\varepsilon) = \inf_{c(x)=\varepsilon} f(x)$	$\sup_y g(y) - \varepsilon^T y$

- x is the primal variable, ε a parameter
- $p^\star(\varepsilon)$ the optimal value when ε varies

Global interpretation of Lagrange multipliers

Hyp: strong duality for the undisturbed problem, that is $y^\star$ t.q.
$g(y^\star) = d^\star = p^\star(0)$
For the perturbated problem we have

$$
\begin{aligned}
p^\star(\varepsilon) &\geq \max_y g(y) - \varepsilon^T y \\
&\geq g(y^\star) - \varepsilon^T y^\star \\
&\geq p^\star(0) - \varepsilon^T y^\star
\end{aligned}
$$

d'où

- if $y_i^\star > 0$ and large, $p^\star$ increases a lot if $\varepsilon_i < 0$
- if $y_i^\star < 0$ and large, $p^\star$ disminishes a lot if $\varepsilon_i > 0$

Local interpretation of Lagrange multipliers

$$p^*(\varepsilon) = \inf_{C(x)=\varepsilon} f(x) \qquad C : \mathbb{R}^n \to \mathbb{R}^m$$

$$y_i^\star = -\frac{\partial p^\star(0)}{\partial \varepsilon_i} \quad \text{for } i = 1, \ldots, m$$

Proof : $\varepsilon = te_i$ in the global sensitivity
$p^\star(te_i) \geq p^\star(0) - ty_i^\star$

$$\lim_{t \searrow 0} \frac{p^\star(te_i) - p^\star(0)}{t} \geq -y_i^\star$$

$$\lim_{t \nearrow 0} \frac{p^\star(te_i) - p^\star(0)}{t} \leq -y_i^\star$$

Exemple 3 : Diagonalization of a symetric matrix

$\ell(x,y) = \langle Ax, x\rangle + y(\|x\|^2 - 1) \qquad y \in \mathbb{R}$

$\nabla_x \ell(x,y) = 2Ax + 2yx$

$$\inf_{\|x\|=1} \langle Ax, x\rangle$$

with A a symetric matrix in $\mathbb{R}^{n\times n}$. $\quad A \in S^n$

$$\inf_{c(x)=0} f(x) \quad \text{with } f(x) = \langle Ax, x\rangle \text{ and } c(x) = \|x\|^2 - 1$$

$Jc(x) = 2x^T \qquad \text{Rank } Jc(x^*) = 1$ because $x^* \neq 0$

- Existence of a minimum since f is continuous and $\{x, \|x\| = 1\}$ bounded closed set.
- f differentiable and $\{c(x) = 0\}$ Lagrange multipliers $\Rightarrow \exists y^\star \in \mathbb{R}$ s.t. $2Ax^\star + 2y^\star x^\star = 0$ $\qquad v = x^* \quad \lambda = -y^*$

  $\Rightarrow \exists(\lambda, v) \in \mathbb{R} \times \mathbb{R}^n,\ Av = \lambda v$ and $f(v) = \inf_{\|x\|=1} f(x)$.

  Recurrence hypothesis H_n : existence of a orthonormal eigenvector basis of A with n related eigenvalues

  $n = 1$ easy

Suppose H_n true

Apply the theorem $\Rightarrow$ $Ax^* = \lambda^* x^*$
$= -y^* x^*$

For $A \in \mathbb{R}^{n+1 \times n+1}$ we consider the subspace $H = \{\text{vect}(x^\star)\}^\perp$.
dim $H = n$
H is stable by A. Indeed $\quad x \in H : \langle x, x^* \rangle = 0$

$$\text{if } \langle x^\star, x \rangle = 0 \text{ then } \langle x^\star, Ax \rangle = \langle Ax^\star, x \rangle = \langle -y^\star x^\star, x \rangle = 0$$

The restriction of A to H is a matrix $n \times n$ therefore using H_n existence of a orthonormal eigenvector basis of the restriction of A to H.
We divide $x^\star$ by $\|x^\star\|$ in order to complete this basis on $\mathbb{R}^{n+1}$.

Exemple 4 : Minimization of a quadratic function under linear constraints of equality

$\inf_{c(x)=0} f(x)$ $\quad c : \mathbb{R}^n \to \mathbb{R}^m$

$$
\begin{aligned}
f(x) &= \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle \\
c(x) &= Bx - C
\end{aligned}
$$

$A \in S^n_{++}$

with A defined symetric positive matrix in $\mathbb{R}^{n\times n}$, b vector in $\mathbb{R}^n$, B matrix in $\mathbb{R}^{m\times n}$ and C vector in $\mathbb{R}^m$.
Qualified constraints $\Leftrightarrow$ rang$(B) = m$.
Lagrangian :

$$
\ell(x, y) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + \langle y, Bx - C\rangle
$$

Theorem of Lagrange multipliers

$\ell(x,y) = \frac{1}{2}\langle Ax, x\rangle + \langle b, x\rangle + \langle y, Bx - C\rangle$

there should be stars everywhere below:

$\underbrace{\text{Rank } B}_{m} + \text{dim Ker } B = \text{dim } n$

$\underbrace{\text{Rank } B^T}_{m} + \underbrace{\text{dim Ker } B^T}_{\overset{\shortparallel}{0}} = m$

$$\nabla_x \ell(x,y) = \quad Ax + b + B^t y = 0$$
$$Bx = C$$

A defined symetric positive matrix $\Rightarrow x = -A^{-1}(b + B^t y)$.
$B(-A^{-1}(b + B^t y)) = C$
$\text{rang}(B) = m \Rightarrow \underline{BA^{-1}B^t \text{ is invertible}}$
$BA^{-1}B^t y = -(BA^{-1}b + C)$ from which we get y then x.

proof

$\hookrightarrow \langle BA^{-1}B^t y, y\rangle = \langle A^{-1}B^t y, B^t y\rangle = 0 \Rightarrow B^t y = 0$

$\text{rang } B = m \Rightarrow y = 0$

$y = -(BA^{-1}B^t)^{-1}(BA^{-1}b + C)$

$x = -A^{-1}\left(b - B^t(BA^{-1}B^t)^{-1}(BA^{-1}b + C)\right)$

Utilisation : SQP algorithm -linear equality constraints

Let the minimization problem with linear equality constraints

$$\left\{\begin{array}{ll}\inf & f(x)\\ s.c. & Bx - c = 0\\ & x \in \mathbb{R}^n\end{array}\right. \quad \text{with} \quad \left\{\begin{array}{l} f : \mathbb{R}^n \longrightarrow \mathbb{R},\ \text{twice differentiable}\\ B \in \mathcal{M}_{m\times n}(\mathbb{R}),\\ c \in \mathbb{R}^m.\end{array}\right.$$

The Lagrangian is

$$\ell(x,y) = f(x) + \langle y, Bx - c\rangle$$

The 1st order optimality constrainst are

$$\begin{array}{rcl}\nabla_x \ell(x,y) & = & \nabla f(x) + B^T y = 0_{\mathbb{R}^n}\\ \nabla_y \ell(x,y) & = & Bx - c = 0_{\mathbb{R}^m}\end{array}$$

Let $G = \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$, $G(x,y) = \begin{pmatrix}\nabla f(x) + B^T y\\ Bx - c\end{pmatrix}$ and use the Newton method to find its zero in $\mathbb{R}^{n+m}$

SQP algorithm : Newton method in $\mathbb{R}^{n+m}$

$G : \mathbb{R}^{n+m} \to \mathbb{R}^{n+m}$

$$G(x,y) = \begin{array}{c} n \\ m \end{array} \begin{pmatrix} \nabla f(x) + B^T y \\ Bx - c \end{pmatrix} \qquad JG(x,y) = \begin{pmatrix} Hf(x) & B^T \\ B & 0_{m\times m} \end{pmatrix}$$

Newton method : $\begin{cases} x_{k+1} = x_k + d_k \\ y_{k+1} = y_k + \delta_k \end{cases}$

$$JG(x_k, y_k) \begin{pmatrix} d_k \\ \delta_k \end{pmatrix} = -G(x_k, y_k)$$

Leads to

$$\begin{cases} Hf(x_k)d_k + \nabla f(x_k) + B^T y_k + B^T \delta_k = 0 = \nabla \tilde{f}(d) \\ Bd_k = 0 \end{cases}$$

Which is equivalent to solve

$$\inf_{Bd=0} \frac{1}{2}\langle Hf(x_k)d, d\rangle + \langle \nabla f(x_k) + B^T y_k, d\rangle = \tilde{f}(d)$$

Algorithme SQP - linear equality constraints

the Lagrange multiplier

At each iteration k , I know x_k, y_k

- let $J_k(d) = \frac{1}{2}\langle Hf(x_k)d, d\rangle + \langle \nabla f(x_k) + B^T y_k, d\rangle$
- minimize $J_k(d)$ under constraint $C(d) = Bd = 0$ use slide number 249

$$\ell(d, \delta) = J_k(d) + \langle \delta, Bd\rangle$$

$\nabla_d \ell(d, \delta) = 0$ and $Cd = 0$ leads to the linear system

$$\begin{pmatrix} Hf(x^k) & B^T \\ B & 0 \end{pmatrix} \begin{pmatrix} d \\ \delta \end{pmatrix} = \begin{pmatrix} -\nabla f(x^k) - B^T y_k \\ 0_{\mathbb{R}^m} \end{pmatrix}$$

$x_{k+1} = x_k + d$

$y_{k+1} = y_k + \delta$

$\begin{cases} x_k + \alpha d \\ y_k + \alpha\delta \end{cases}$ instead

Algorithme SQP - linear equality constraints

3 pitfalls

1. $Hf(x^k)$ may be hard to compute : quasi-Newton approximation $\hat{H}$
2. $\hat{H}$ may be not invertible : penalize with $\max(0, -\min(\lambda_{\hat{H}}) + \varepsilon)\, Id$
3. $\nabla_\delta \ell(\delta, y)$ might not decrease -> line search for a better step

Algorithme SQP - linear equality constraints

Data: Function f, gradient ∇f, hessien Hf, tolerance τ, max number of iterations $k_{\max}$

Result: $f(x^\star) = \min_{x \in \mathbb{R}^n} f(x)$

Initialisation : choose $x_0 \in \mathbb{R}^n$, $d_0 \in \mathbb{R}^n$ t.q. $\|d_0\| > \tau$

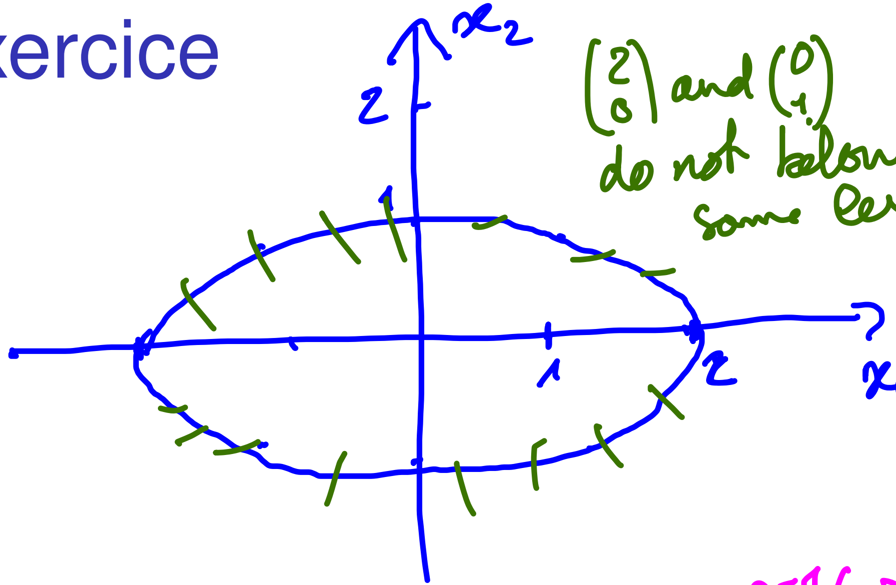
while $\|d^k\| \geq \tau$ *and* $k < k_{\max}$ **do**

- Compute $f(x^k)$, $\nabla f(x^k)$ et $Hf(x^k)$
- Minimize $J_k(d) = \frac{1}{2}\langle Hf(x_k)d, d\rangle + \langle \nabla f(x_k) + B^T y_k, d\rangle$ under constraints $Bd = 0 \quad \rightarrow$ find $d^\star$ and $\delta^\star$
- Update $d_k = d^\star$,
- Update $x^{k+1} = x^k + d^\star, \quad y^{k+1} = y^k + \delta^\star$
- Update $k \leftarrow k + 1$

end

$x^\star = x_k$

Exercice



$\binom{2}{0}$ and $\binom{0}{1}$ do not belong to the same level curve

$f\binom{2}{0} = 4$

$f\binom{0}{1} = 2$

$f(x) = x_1^2 + 2x_2^2$

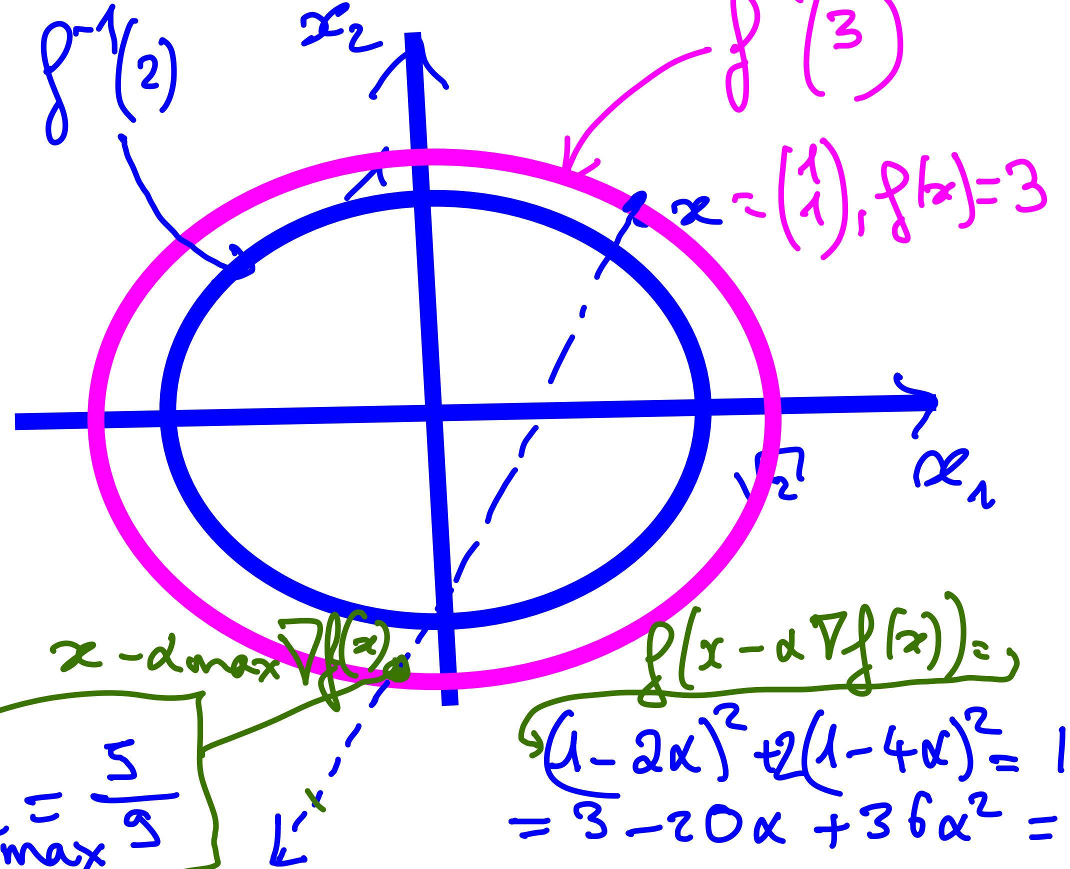
1) which level curve is drawn?

2) set $x = \binom{1}{1}$

compute $-\nabla f(x)$

set $x - \nabla f(x)$ on the graph and draw the line

3) compute α_{max} such that $f(x - \alpha \nabla f(x)) < f(x)$ for $0 < \alpha < \alpha_{max}$



$-\nabla f(x) = -\binom{2x_1}{4x_2} = -\binom{2}{4}$

$f(x - \alpha \nabla f(x)) =$

$(1-2\alpha)^2 + 2(1-4\alpha)^2 = 1 - 4\alpha + 4\alpha^2 + 2 - 16\alpha + 32\alpha^2$

$= 3 - 20\alpha + 36\alpha^2 = 3$ when $-5 + 9\alpha = 0$

$\boxed{\alpha_{max} = \frac{5}{9}}$