Exercice

$$f(x) = 3x_1^2 + x_2^2 - 2x_1x_2 + x_1 + x_2 + 1$$

write in the quadratic form A, b, c

Correction:

$$\nabla f(x) = \begin{pmatrix} 6x_1 - 2x_2 + 1 \\ 2x_2 - 2x_1 + 1 \end{pmatrix} \qquad Hf(x) = \begin{pmatrix} 6 & -2 \\ -2 & 2 \end{pmatrix} = \text{cte} = A$$

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

$$= \underset{c}{\underset{\parallel}{1}} + \langle x, \underset{b}{\underset{\parallel}{\begin{pmatrix} 1 \\ 1 \end{pmatrix}}} \rangle + \frac{1}{2}\langle Ax, x \rangle$$

$$\det(A - \lambda I) = (6-\lambda)(2-\lambda) - 4 = \lambda^2 - 8\lambda + 8 =$$

$$\Delta = 64 - 32 = 32 = (\sqrt{2}\,4)^2$$

$$\lambda = \frac{8 \pm 4\sqrt{2}}{2} > 0$$

$A \in S^n_{++} \Rightarrow Ax^* + b = 0$ has a unique solution which is the global minimum of f: $\left(-\frac{1}{2}, -1\right) = x^*$

Exercice

$f : \mathbb{R}^2 \to \mathbb{R}$

$f(x_1, x_2) = 4(x_1^2 + x_2^2) - (x_1^2 + x_2^2)^2 \qquad X = (x_1, x_2)$

1) Compute $\nabla f(X)$ and $Hf(X)$

2) Compute the set of points $\{\nabla f(X) = 0\} = S$

3) Say if $X \in S$ is minimum or maximum or what?

$$\nabla f(X) = \begin{pmatrix} 8x_1 - 4x_1(x_1^2 + x_2^2) \\ 8x_2 - 4x_2(x_1^2 + x_2^2) \end{pmatrix} \qquad Hf(x) = \begin{pmatrix} 8 - 12x_1^2 - 4x_2^2 & -8x_1x_2 \\ -8x_1x_2 & 8 - 12x_2^2 - 4x_1^2 \end{pmatrix}$$

$$\begin{pmatrix} 2x_1 = x_1(x_1^2 + x_2^2) \\ 2x_2 = x_2(x_1^2 + x_2^2) \end{pmatrix} \iff S = \{(0,0)\} \cup \left\{ \begin{array}{l} x_1^2 + x_2^2 = 2 \\ \text{circle of center } (0,0) \\ \text{radius } \sqrt{2} \end{array} \right\}$$

question 3 for tomorrow

$z = \|x\|^2, \quad f(x) = 4z - z^2$

Outline

Canonical problem

$$
\begin{cases}
\text{inf} & f(x) \\
s.c. & c^E x) = 0 \\
s.c. & c^I(x) \leq 0 \\
 & x \in \mathbb{R}^n
\end{cases}
$$

equality constraints
inequality constraints

with

$$
\begin{array}{rcl}
f & : & \mathbb{R}^n \longrightarrow \mathbb{R}, \\
c^E & : & \mathbb{R}^n \longrightarrow \mathbb{R}^m, \\
c^I & : & \mathbb{R}^n \longrightarrow \mathbb{R}^p, \\
f & c & \text{smooth.}
\end{array}
$$

m eq. constraints
p ineq. const.

admissible domain

$$
D_a = \{x \in \mathbb{R}^n,\ c^E(x) = 0,\ c^I(x) \leq 0\}
$$

General existence theorem

$$\inf_{x \in D_a} f(x)$$

if $D_a = C$, f continuous

We consider f continuous from $C \subset \mathbb{R}^n$ into $\mathbb{R}$ with C closed. If one of the following hypotheses is satisfied

- C bounded
- C not bounded and f coercive

$$f(x) \longrightarrow +\infty$$
$$\|x\| \longrightarrow +\infty$$

example $f(x) = \|x\|^2$

then f has a minimum on C

$c^E : \mathbb{R}^n \longrightarrow \mathbb{R}^m$

$c^I : \mathbb{R}^n \longrightarrow \mathbb{R}^p$

$c^E = (c_i^E)_{i=1,\dots m}$ $c_i^E : \mathbb{R}^n \longrightarrow \mathbb{R}$

Notations for the gradient and the hessian of the i^{th} constraint

$$a_i^E(x) = \nabla c_i^E(x) \quad H_i^E(x) = Hess\ c_i^E(x),$$
$$a_i^I(x) = \nabla c_i^I(x) \quad H_i^I(x) = Hess\ c_i^I(x).$$

Jacobian Matrices of the constraints :

$$A^E(x) = \nabla c^E(x) = \begin{pmatrix} a_1^E(x)^T \\ \vdots \\ a_m^E(x)^T \end{pmatrix}, \quad A^I(x) = \nabla c^I(x) = \begin{pmatrix} a_1^I(x)^T \\ \vdots \\ a_p^I(x)^T \end{pmatrix}$$

Lagrangian and Lagrange multipliers

$\left.\begin{array}{l}\inf f(x) \\ c^E(x) = 0 \\ c^I(x) \le 0\end{array}\right) (P)$

(P) direct or primal problem

Let y a vector of $\mathbb{R}^m$, z a vector of $\mathbb{R}^p$, Lagrange multipliers.

dual variable

The Lagrangien is defined by

$\ell : \mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$

$$\ell(x,y,z) = f(x) + \langle y, c^E(x)\rangle + \langle z, c^I(x)\rangle$$

$+ \sum_{i=1}^{m} y_i c_i^E(x) + \sum_{j=1}^{p} z_j c_j^I(x)$

The gradient and the hessian of the Lagrangienwith respect to x are

$\sum y_i \nabla c_i^E(x) + \sum z_j \nabla c_j^I(x)$

$$g(x,y,z) = \nabla_x \ell(x,y,z) = \nabla f(x) + \sum_{i=1}^{m} y_i a_i^E(x) + \sum_{i=1}^{p} z_i a_i^I(x)$$

$$H(x,y,z) = Hess_x \ell(x,y,z) = Hf(x) + \sum_{i=1}^{m} y_i H_i^E(x) + \sum_{i=1}^{p} z_i H_i^I(x)$$

Example 1: $f : \mathbb{R}^2 \to \mathbb{R} \quad f(x) = x_1 + x_2, \quad \inf_{x_1^2 + x^2 = 2} f(x)$

$c^E : \mathbb{R}^2 \to \mathbb{R} \qquad m = 1, \; p = 0$

$c^E(x) = x_1^2 + x_2^2 - 2 \qquad c^E(x) = 0$

$\ell : \mathbb{R}^2 \times \mathbb{R} \to \mathbb{R}$

$\ell(x, y) = f(x) + y\, c^E(x) = x_1 + x_2 + y(x_1^2 + x_2^2 - 2)$

$$\nabla_x \ell(x, y) = \begin{pmatrix} 1 \\ 1 \end{pmatrix} + y \begin{pmatrix} 2x_1 \\ 2x_2 \end{pmatrix} = \begin{pmatrix} 1 + 2yx_1 \\ 1 + 2yx_2 \end{pmatrix}$$

$$H_x \ell(x, y) \in \mathcal{M}_{2\times 2}(\mathbb{R}); \; H_x \ell(x, y) = \begin{pmatrix} 2y & 0 \\ 0 & 2y \end{pmatrix}$$

only inequality constraints

Example 2: $f : \mathbb{R}^n \to \mathbb{R}$ $\quad f(x) = ||x||^2, \quad \inf_{\substack{x_{i+1} - x_i \le 2 \\ i = 1, \dots, n-1}} f(x)$

$m = 0, \quad p = n-1$

$l : \mathbb{R}^n \times \mathbb{R}^p \to \mathbb{R}$ $\qquad c^I : \mathbb{R}^n \to \mathbb{R}^p = \mathbb{R}^{n-1}$

$l(x, z) = f(x) + \langle z, c^I(x) \rangle$ $\qquad x \longmapsto \begin{pmatrix} x_2 - x_1 - 2 \\ \vdots \\ x_n - x_{n-1} - 2 \end{pmatrix}$

$= ||x||^2 + \sum_{j=1}^{n-1} z_j (x_{j+1} - x_j - 2) = ||x||^2$

$+ z_1 (x_2 - x_1 - 2)$

$+ z_2 (x_3 - x_2 - 2)$

$\vdots$

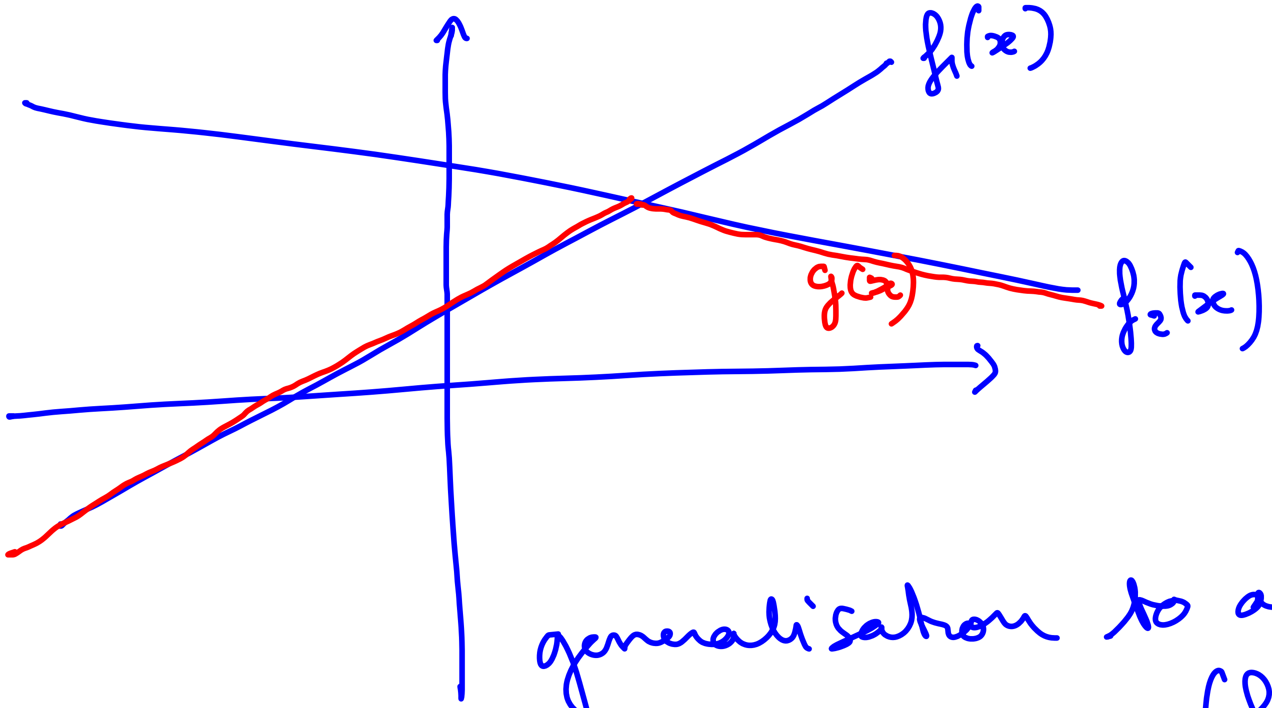
$+ z_{n-2} (x_{n-1} - x_{n-2} - 2)$

$+ z_{n-1} (x_n - x_{n-1} - 2)$

$$\nabla_x l(x, z) = 2x + \begin{pmatrix} -z_1 \\ +z_1 - z_2 \\ \vdots \\ +z_{n-2} - z_{n-1} \\ +z_{n-1} \end{pmatrix}$$

$$H_x l(x, z) = 2 I_{d_{n \times n}} + 0$$

[redacted]



each f_i is concave

$$f(\lambda x+(1-\lambda)y) \geq \lambda f(x) + (1-\lambda) f(y)$$

$$g(x) = \min_{i=1,2} f_i(x)$$

g is concave

generalisation to any family of $(f_i(x))_{i\in I}$ f_i concave fxn

generalize to any uncountable family $(f(x,p))_{p\in C}$ when $f(x,p)$ is concave for each p

then $g(x) = \min_{p\in C} f(x,p)$ is concave

Actives Constraints / Contraintes actives in the case of inequality constraints

Let $x^\star$ a minimizer of f.

The i^{th} inequality constraint is active if $c_i^I(x^\star) = 0$.

$\inf f(x) = f(x^*)$ $\quad x^* \in \mathcal{D}_a = \{ c^E(x)=0,\ c^I(x) \leq 0 \}$

$c^I(x) \leq 0$
and, possibly, $c^E(x) = 0$ $\Rightarrow c^E(x^*) = 0 \quad \underbrace{c^I(x^*) \leq 0}$

$c_j^I(x^*) \leq 0 \quad j = 1, \dots, p$

$c_j^I(x^*) = 0$ or $c_j^I(x^*) < 0 \quad j = 1, \dots, p$

$\Downarrow \qquad\qquad \Downarrow$

c_j^I active at x^* $\qquad$ c_j^I is inactive at x^*

Actives Constraints / Contraintes actives

Let $x^\star$ a minimizer of f.

The i^{th} inequality constraint is active if $c_i^I(x^\star) = 0$.

1. $f : \mathbb{R}^2 \to \mathbb{R} \quad f(x) = ||x||^2, \quad \inf_{x_1+x_2 \leq 1} f(x)$

$f^{-1}(0) = \{(0,0)\}$

$f^{-1}(1) =$ circle

$c^I : \mathbb{R}^2 \longrightarrow \mathbb{R}$

$x \longmapsto x_1 + x_2 - 1$

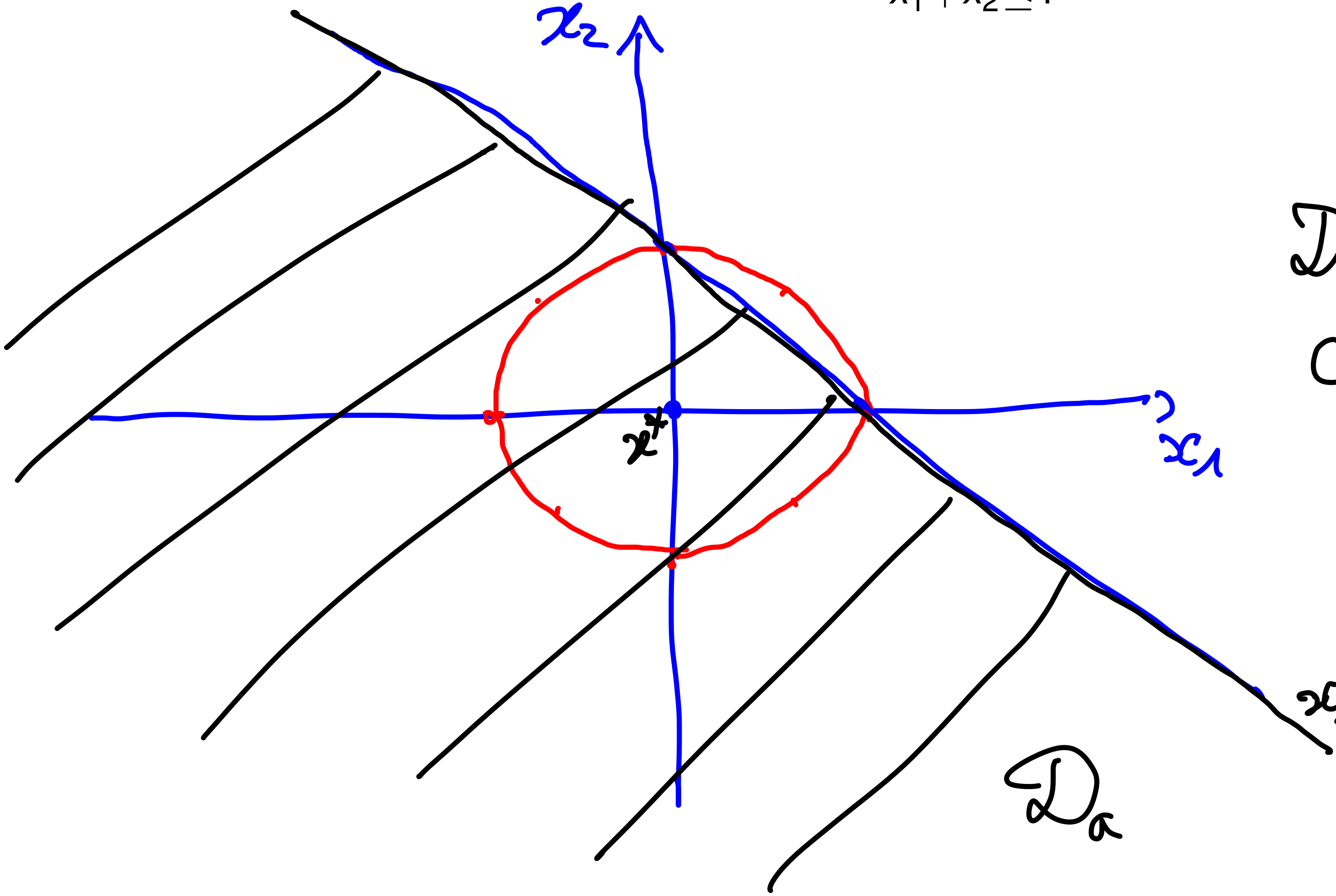
$x^\star = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \in \mathcal{D}_a$

$\mathcal{D}_a = \{x \in \mathbb{R}^2, \ c^I(x) \leq 0\}$

c^I active or not?

$c^I(x^\star) = -1$

inactive



$x_1 + x_2 = 1$

In green : second example

Actives Constraints / Contraintes actives

$$c^I : \mathbb{R}^2 \to \mathbb{R}^2, \quad x \mapsto \begin{pmatrix} x_1 + x_2 + 1 \\ x_2 \end{pmatrix}$$

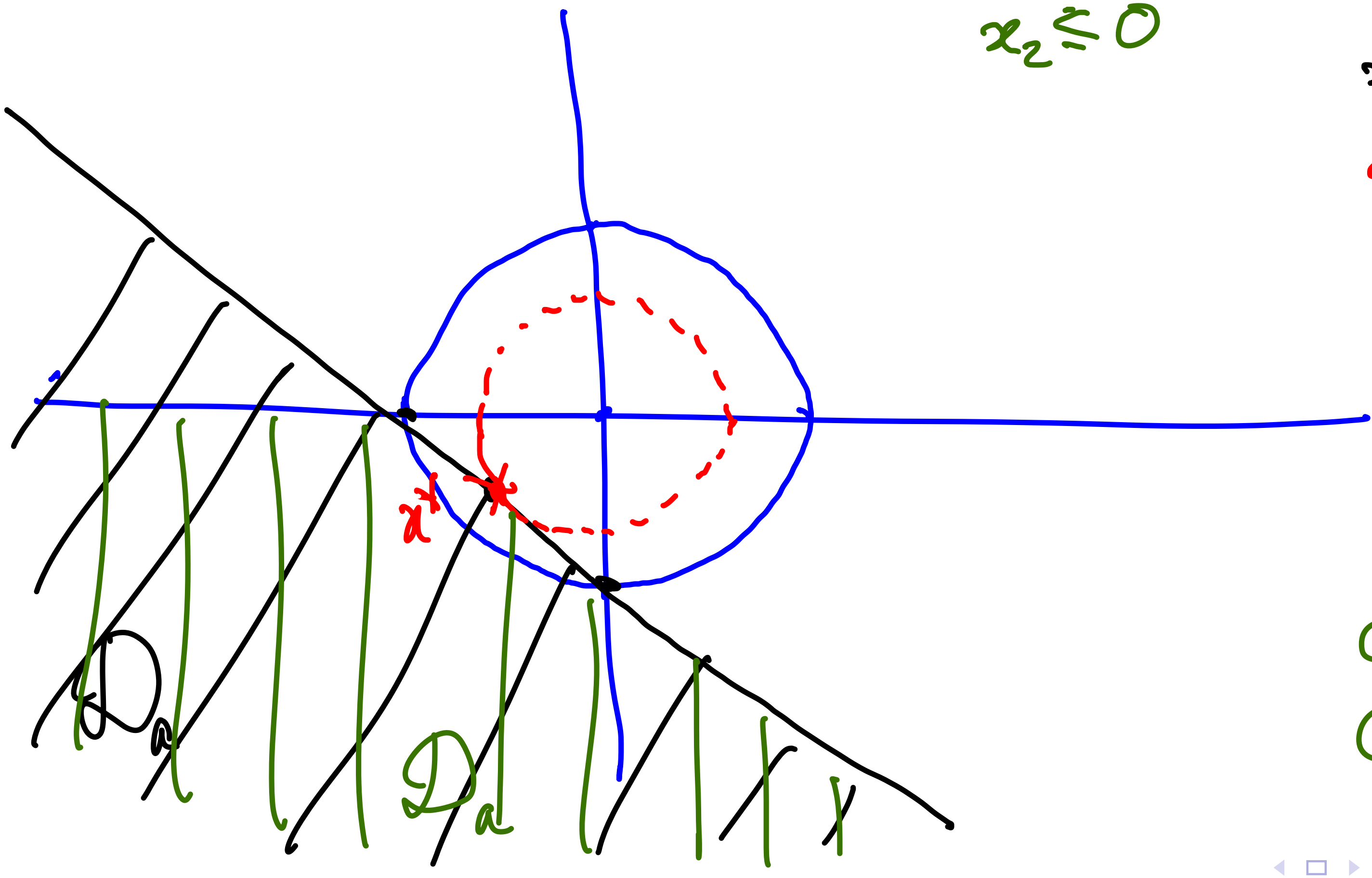
Let $x^\star$ a minimizer of f.

The i^{th} inequality constraint is active if $c_i^I(x^\star) = 0$.

2. $f : \mathbb{R}^2 \to \mathbb{R} \quad f(x) = ||x||^2, \quad \inf_{x_1 + x_2 \leq -1} f(x)$

$x_2 \leq 0$

$$c^I : \mathbb{R}^2 \to \mathbb{R}, \quad x \mapsto x_1 + x_2 + 1$$



$x^* \in \mathcal{D}_a$

$x^* = \begin{pmatrix} -1/2 \\ -1/2 \end{pmatrix}$

$c^I(x^*) = 0$

therefore c^I is active at x^*

c_1^I is active

c_2^I is not active

Lagrange dual function

x primal variable

$\inf f(x)$ —— problem

y, z —— dual variable

proof: $\ell(x^*, y, z) = f(x^*) + \langle z, c^I(x^*)\rangle \leq p^*$

$f(x^*) = p^*$, $c^I(x^*) \leq 0$

$$g : \mathbb{R}^m \times \mathbb{R}^p \to \mathbb{R}$$

$$\begin{aligned} g(y,z) &= \inf_{x \in D_a} \ell(x,y,z) \leq \ell(x^*, y, z) \leq p^* \\ &\quad \left(f(x) + \langle y, c^E(x)\rangle + \langle z, c^I(x)\rangle\right) \\ &= \inf_{x \in D_a} \left(f(x) + \sum_{i=1}^{m} y_i c_i^E(x) + \sum_{i=1}^{p} z_i c_i^I(x) \right) \end{aligned}$$

g is concave (can be unbounded for some y,z)

Property : inferior bound:

If $z \geq 0$ then $g(y,z) \leq p^\star = \inf_{x \in D_a} f(x) = f(x^*)$

-> an affine function is both convex and concave

-> for ~~fi~~ fixed x $\quad (y, z) \longmapsto f(x) + \langle y, c^E(x)\rangle + \langle z, c^I(x)\rangle$ concave in y, z

Exemple : solution of a linear system with minimal norm

$A \in \mathcal{M}_{m\times n}(\mathbb{R})$ $b \in \mathbb{R}^m$

Solve

$$p^\star = \inf_{Ax=b} x^T x = \|x\|^2$$

Primal problem

$c^E : \mathbb{R}^n \to \mathbb{R}^m$

$c^E(x) = Ax - b$

- Lagrangian : $\ell(x,y) = x^T x + y^T(Ax-b)$

$= \|x\|^2 + \langle y, c^E(x)\rangle$ coercive

~~fix y and look at $\inf \ell(x,y)$~~ does it exists?

$g(y) = \inf_{x \in D_a} \ell(x,y)$

$\nabla_x \ell(x,y) = 2x + A^T y$ $\quad x = -\frac{1}{2}A^T y$

$g(y) = \frac{1}{4}\langle A^T y, A^T y\rangle + \langle y, -\frac{1}{2}AA^T y - b\rangle = -\frac{1}{4}\|A^T y\|^2 - \langle y, b\rangle$

$g(y) \leq p^*$ $\forall y$ moreover

$g(y)$ is concave we can compute its maximum value

Exemple : solution of a linear system with minimal norm

Solve

$$p^\star = \inf_{Ax=b} x^T x$$

- Lagrangian : $\ell(x, y) = x^T x + y^T (Ax - b)$
- In order to minimize $\ell(x, y)$ with respect to x we seek gradient zeros

$$\nabla_x \ell(x, y) = 2x + A^T y = 0 \;\Leftarrow\; x = -A^T y/2$$

Exemple : solution of a linear system with minimal norm

Solve

$$p^\star = \inf_{Ax=b} x^T x$$

- Lagrangian : $\ell(x, y) = x^T x + y^T (Ax - b)$
- In order to minimize $\ell(x, y)$ with respect to x we seek gradient zeros

$$\nabla_x \ell(x, y) = 2x + A^T y = 0 \Leftarrow x = -A^T y/2$$

- Inject in the definition of the dual function

$$g(y) = \ell(-A^T y/2, y) = -\frac{1}{4} y^T A A^T y - b^T y$$

concave in y

Exemple : solution of a linear system with minimal norm

Solve

$$p^\star = \inf_{Ax=b} x^T x$$

- Lagrangian : $\ell(x, y) = x^T x + y^T(Ax - b)$
- In order to minimize $\ell(x, y)$ with respect to x we seek gradient zeros

$$\nabla_x \ell(x, y) = 2x + A^T y = 0 \;\Leftarrow\; x = -A^T y/2$$

- Inject in the definition of the dual function

$$g(y) = \ell(-A^T y/2, y) = -\frac{1}{4} y^T A A^T y - b^T y$$

concave in y

- Inferior bound property

$$p^\star \geq -\frac{1}{4} y^T A A^T y - b^T y \;\forall y$$

Resolution of the dual problem

since $g(y,z) \le p^*$ for all y, z

$d^* = \sup_{y,z \geq 0} g(y,z) \leq p^*$

Solve

$$d^\star = \sup_{y \in \mathbb{R}^m, z \in \mathbb{R}^p, z \geq 0} g(y,z)$$

- Best inferior bound for $p^\star \geq d^\star$
- The dual problem is concave : existence of an optimal problem $d^\star$

Weak duality $d^\star \leq p^\star$

Strong duality $d^\star = p^\star$

many hypothesis on strong duality

Slater qualification of constraints ensure strong duality

Qualified constraints

Slater hypothesis

- Condition for $d^\star = p^\star$
- We say that the constraints are qualified in $x^\star$ if the rank of the matrix formed by the union of the Jacobian matrix of equality constraints and the Jacobian matrix of q constraints of active inequality in $x^\star$ is equal to $m + q$, then called maximal rank.
- Particular case of a convex problem :
  Solve

  f ~~convex~~ c^I, convex

  $$p^\star = \inf_{\substack{Ax = b \\ c^I(x) \leq 0}} f(x)$$

  if there exist $x \in \mathring{\mathcal{D}}_a$

If $\exists x$ s.t. $c^I(x) < 0$ and $Ax = b$ then $d^\star = p^\star$

Case of equality constraints

C qualified $\Longleftrightarrow$ Rank $JC(x^*) = m$ or C affine

$$\begin{cases} \inf & f(x) \\ s.c. & C(x) = 0 \\ & x \in \mathbb{R}^n \end{cases}$$

with

$$\begin{array}{lll} f & : & \mathbb{R}^n \longrightarrow \mathbb{R}, \\ C & : & \mathbb{R}^n \longrightarrow \mathbb{R}^m, \\ f & c & \text{smooth.} \end{array}$$

Théorème des extrema liés - Lagrange multipliers

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

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

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

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

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