Optimisation linéaire

(I) Définition du problème

Un problème d'optimisation linéaire est un problème d'optimisation dans lequel on minimise une fonction linéaire sur un polyèdre convexe. Si $(E, \langle \cdot, \cdot \rangle)$ est l'espace euclidien sous-jacent, il s'écrit

$$(P_c) \quad \inf_{x \in P} \langle c, x \rangle,$$

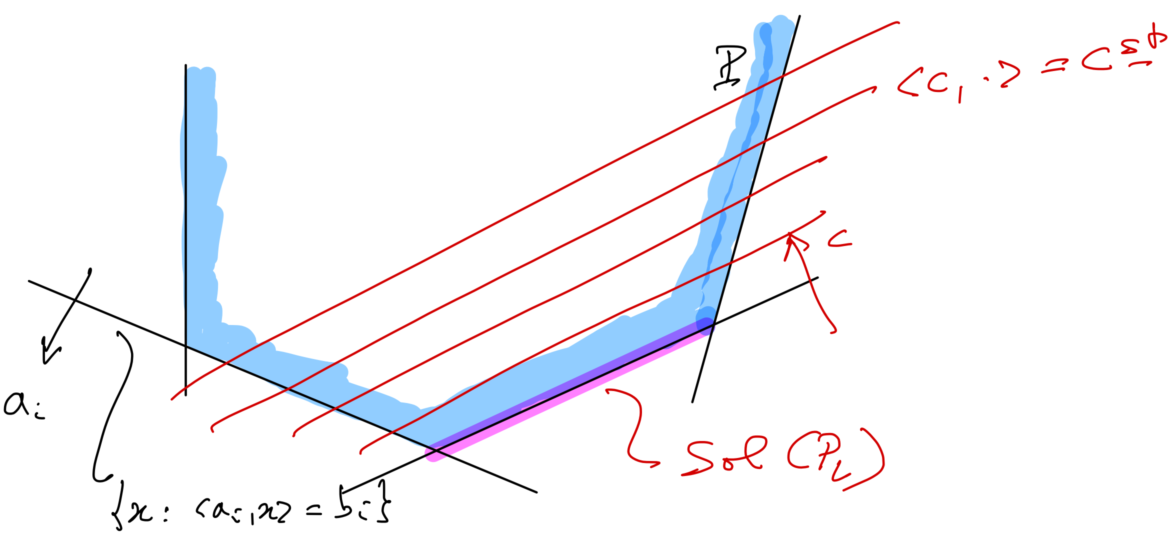
où $c \in E$ et P est un polyèdre convexe.

Il y a plusieurs manières de représenter P

Forme canonique

$P = \bigcap$ finie de demi-espaces fermés

Cela permet de faire des dessins



$$P = \{x \in E : Ax \leq b\}$$

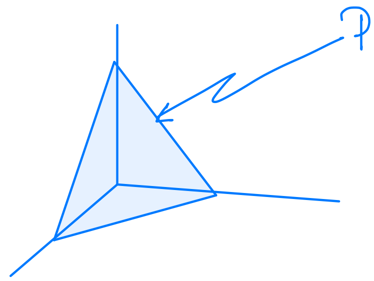
$\langle a_i, x \rangle \leq b_i, \ \forall i \in [1:m]$

Forme standard (dans $\mathbb{R}^n$)

$P = \cap$ de l'orthant positif $\mathbb{R}^n_+$ et d'un sous-espace affine

$$= \{ x \in \mathbb{R}^n : Ax = b,\ x \geq 0 \}$$

Cela permet de généraliser ($\mathbb{R}^n_+ \curvearrowright$ Cone)



Les 2 formes sont "équivalentes"

- Canonique $\curvearrowright$ standard

$Ax \leq b \longrightarrow Ax + s = b,\ s \geq 0$ (variables d'écart)

$x = u - v,\ u \geq 0,\ v \geq 0$

$$\longrightarrow \begin{cases} Au - Av + s = b \\ (u, v, s) \geq 0 \end{cases}$$

- Standard $\curvearrowright$ Canonique

$$\begin{matrix} Ax = b \\ x \geq 0 \end{matrix} \longrightarrow \begin{cases} Ax \leq b \\ -Ax \leq -b \\ -x \leq 0 \end{cases}$$

$Ax = b$

$\ell \leq Bx \leq u$

$\ell_x \leq x \leq u_x$

Buts du cours

① **étudier le pbl** (tout a déjà été vu !!)

- $\exists$ de sol
- CO
- dualité

② **Algorithmique**

- comparaison de l'algo du simplexe (ou de points-frontière) et les algos de points-intérieurs
- comment obtient-on la polynomialité par les PI

II Etude du problème

1 Existence de solution

$$(P_c)\begin{cases} \inf \langle c, x\rangle \\ x \in P \end{cases} \qquad \begin{array}{l} c \in E \\ P = \text{polyèdre convexe} \end{array}$$

(P_c) a une solution

$\Longleftrightarrow$ $\text{val}(P_c) \in \mathbb{R}$

$\Longleftrightarrow$ $\begin{cases} (P_c) \text{ est réalisable} & (\text{val}(P_c) < +\infty) \\ (P_c) \text{ est borné} & (\text{val}(P_c) > -\infty) \end{cases}$

Dém $\Rightarrow$ évident

$\Leftarrow$ $I = \{ \langle c, x\rangle : x \in P \} \subset \mathbb{R}$

$\hookrightarrow$ polyèdre convexe $\Longrightarrow$ un intervalle de $\mathbb{R}$ fermé

Soit (x_k) suite minimisante du problème $(P \neq \emptyset)$

$(x_k) \subset P$, $\underbrace{\langle c, x_k\rangle}_{\in I \text{ fermé}} \longrightarrow \text{val}(P_c)$

$\Rightarrow$ $\text{val}(P_c) \in I$

$\Rightarrow$ $\text{val}(P_c) = \langle c, \bar{x}\rangle$, $\bar{x} \in P$

$\Rightarrow$ $\bar{x} \in \text{Sol}(P_c)$

$\square$

② Conditions d'optimalité

Pour $(P_C)\begin{cases} \min\ c^T x \\ Ax = b \leftarrow y \\ x \geqslant 0 \leftarrow s \end{cases}$ dans $\mathbb{R}^n$

- Lagrangien $\ell(x,y,s) = c^T x - y^T(Ax-b) - s^T x$
- les contraintes sont qualifiées, car affines (QC-A)
- le problème est convexe

Donc

$$\boxed{\begin{array}{l} x \in \mathrm{Sol}(P_C) \\ \Longleftrightarrow \exists (y,s) \in \mathbb{R}^m \times \mathbb{R}^n : \\ \qquad \begin{cases} A^T y + s = c \\ Ax = b \\ 0 \leqslant x \perp s \geqslant 0 \end{cases} \end{array}}$$

③ Dualité

$$
(P_L) \begin{cases} \inf\limits_x \ c^T x \\ Ax = b \\ x \geq 0 \end{cases}
$$

$$
\ell(x,y,s) = c^T x - y^T(Ax-b) - s^T x
$$

$$
(P_L) \Leftrightarrow \inf_x \ \sup_{\substack{y \\ s \geq 0}} \ \ell(x,y,s)
$$

$$
(D_L) \Leftrightarrow \sup_{\substack{y \\ s \geq 0}} \ \inf_x \ \ell(x,y,s)
$$

$$
\ell = (c - A^T y - s)^T x + b^T y
$$

$$
(D_L) \begin{cases} \sup \ b^T y \\ A^T y + s = c \\ s \geq 0 \end{cases}
$$

On en déduit la relation de dualité faible

$$
\boxed{\mathrm{val}(D_L) \leq \mathrm{val}(P_L)}
$$

Dualité forte

> Les propriétés suivantes sont équivalentes
> (i) (P_L) et (D_L) sont réalisables
> (ii) (P_L) a une solution
> (iii) (D_L) a une solution
> Dans ces conditions $val(D_L) = val(P_L)$

Dém (rappel)

- (i) $\Rightarrow$ (ii)
  - (P_L) est réalisable $\Rightarrow val(P_L) < +\infty$
  - $\underbrace{val(D_L)}_{> -\infty} \leq val(P_L) \Rightarrow val(P_L) > -\infty$

  (P_L) a 1 solution

- (ii) $\Rightarrow$ (iii) (P_L) a 1 solution $\Rightarrow$ CO : $\exists (y, s)$: $\begin{cases} A^T y + s = c \\ Ax = b \\ 0 \leq x \perp s \geq 0 \end{cases}$

  ce sont les CO de (D_L)

  $\Rightarrow (D_L)$ a 1 solution

- (iii) $\Rightarrow$ (i) (D_L) a 1 solution $\Rightarrow$ CO : $\exists x$: $Ax = b$, $x \geq 0$ $\Rightarrow (P_L)$ réalisable
- $c^T \bar{x} - b^T \bar{y} = (A^T \bar{y} + \bar{s})^T \bar{x} - b^T \bar{y} = \bar{s}^T \bar{x} = 0$ (compl.) $\square$

Rmq 1) La seule manière d'avoir un saut de dualité est d'avoir $val(P_L) = +\infty$, $val(D_L) = -\infty$

exemple : $A = 0$, $b \neq 0$, $c < 0$

$\Rightarrow val(P_L) = +\infty$ $\qquad \Rightarrow val(D_L) = -\infty$

2) On peut avoir $val(D_L) = val(P_L) \in \{-\infty, +\infty\}$

exemples
- $A = 0$, $b = 0$, $c < 0 \Rightarrow val(P_L) = -\infty$
- $A = 0$, $b \neq 0$, $c \geq 0 \Rightarrow val(D_L) = +\infty$

$\square$

CO du dual

$$(y,s) \in \text{Sol}(D_L) \iff \exists x \text{ tel que } \begin{cases} A^T y + s = c \\ Ax = b \\ 0 \leq s \perp x \geq 0 \end{cases}$$

<u>Dém</u> (D_L) est un problème convexe et ses contraintes sont affines, donc ses CO sont nécessaires et suffisantes

Lagrangien :

car pbl de maximisation

$$\ell((y,s),(x,z)) = -b^T y + x^T(A^T y + s - c) - z^T s$$

$$\begin{array}{l} \nabla_y \ell = 0 \Rightarrow \\ \nabla_s \ell = 0 \Rightarrow \end{array} \begin{cases} Ax = b \\ x = z \\ A^T y + s = c \\ 0 \leq x \perp s \geq 0 \end{cases} \iff \begin{cases} A^T y + s = c \\ Ax = b \\ 0 \leq s \perp x \geq 0 \end{cases}$$

$\square$

Structure cartésienne des CO

$$\left.\begin{array}{l} x \in \text{Sol}(P_L) \\ (y,s) \in \text{Sol}(D_L) \end{array}\right\} \iff \begin{cases} A^T y + s = c \\ Ax = b \\ 0 \leq s \perp x \geq 0 \end{cases}$$

<u>Dém</u> $\Leftarrow$ par les CO de (P_L) et (D_L)

$\Rightarrow$ $x \in \text{Sol}(P_L) \Rightarrow \exists (\bar{y},\bar{s}) :$ $\begin{cases} A^T \bar{y} + \bar{s} = c \\ \boxed{Ax = b} \\ 0 \leq \bar{s} \perp \boxed{x \geq 0} \end{cases}$

$(y,s) \in \text{Sol}(D_L) \Rightarrow \exists \bar{x} :$ $\begin{cases} \boxed{A^T y + s = c} \\ A\bar{x} = b \\ \boxed{0 \leq s} \perp \bar{x} \geq 0 \end{cases}$

Il reste à montrer que $\boxed{s \perp x}$

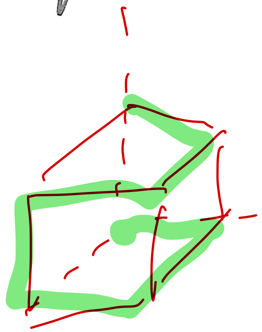
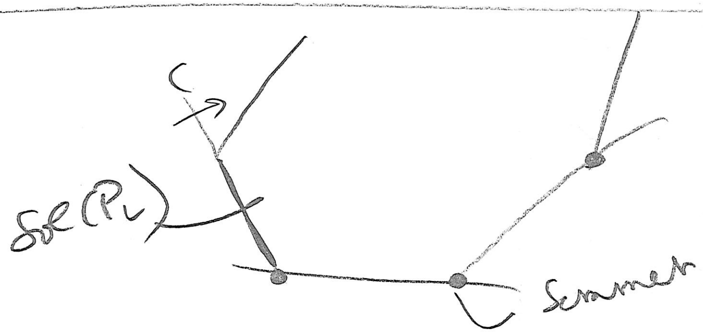
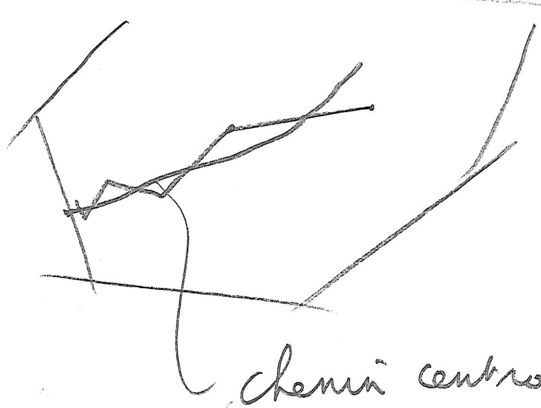
$s^T x = (c - A^T y)^T x = c^T x - b^T y = 0$ car $\text{val}(D_L) = \text{val}(P_L)$

$\square$

III Algorithmique

1 2 familles d'algorithmes

- algos de points-frontière
- ———————— intérieurs

	Simplexe (Points frontière)	Points intérieurs
concept de base 	 l'algo passe de sommet en sommet $c^T x \downarrow$ val(Pc) pour trouver la solution exacte	 l'algo suit ± le chemin central pour trouver la solution à $\varepsilon > 0$ près
Initiateur	DANTZIG (~1947)	KARMARKAR (~1984)
type de convergence	finie	infinie
polynomialité	• non (Klee-Minty, 1972) # iter $= 2^n$ parfois (pire cas) • oui en moyenne (Borgwardt, 1982) (Smale, 1983)	• oui à ε près # iter $= O(n^{1/2} \log \varepsilon^{-1})$

② Polynomialité en PI

A

– Les meilleurs algorithmes sont PD ⇒ s'attaquent à la résolution de

$$\begin{cases} A^T y + s = c, & s \geq 0 \\ Ax = b, & x \geq 0 \\ x^T s = 0 \end{cases}$$

– On prend des points strictement intérieurs c-à-d avec $x > 0$ et $s > 0$

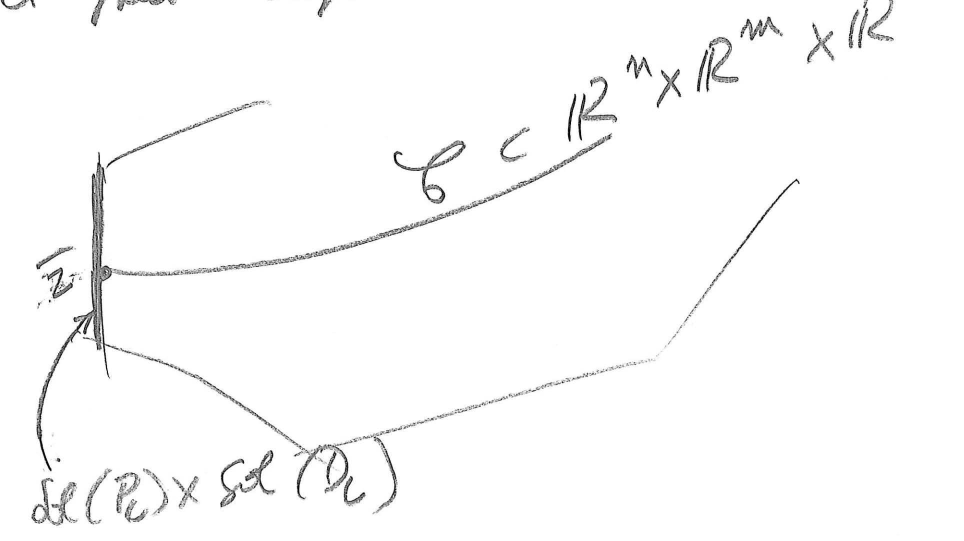
– Nous allons parler des algos "admissibles" c-à-d

$$x \in \mathcal{F}_P^s = \{x \in \mathcal{F}_P : x > 0\}$$

$$(y,s) \in \mathcal{F}_D^s = \{(y,s) \in \mathcal{F}_D : s > 0\}$$

$$\Rightarrow \text{hypothèse : } \mathcal{F}_P^s \times \mathcal{F}_D^s = \boxed{\mathcal{F}^s \neq \emptyset}$$

– Les itérés suivent le chemin central qu'il faut définir



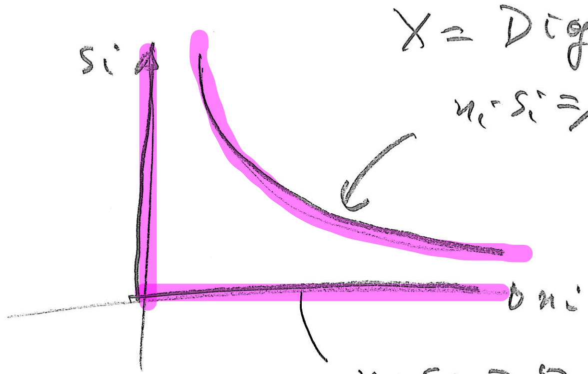
1er point de vue (t-primal)

On remplace dans (Θ)

$x_i s_i = 0, \ \forall i$ par $x_i s_i = \mu, \ \forall i$

$XS = 0$ $XS = \mu e$

$X = \mathrm{Diag}(x_1, \dots, x_n)$ $e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$



$(x_i \geq 0, \ s_i \geq 0)$

Le chemin central $=$ image de l'application

$\mu > 0 \longmapsto z_\mu = (x_\mu, y_\mu, s_\mu)$ solution unique de

$$(\Theta_\mu) \begin{cases} A^T y + s = c & , \ s > 0 \\ Ax = b & , \ x > 0 \\ Xs = \mu e \end{cases}$$

Comment montrer $\exists!$ solution —

retour à l'optimisation (voir ci-après)

2e point de vue (optimisation) on cherche

à éviter le bord de $\mathcal{F}_p$ en pénalisant $x \geq 0$:

$$(P_\mu) \begin{cases} \inf\limits_{x > 0} \ c^T x - \mu \sum\limits_{i=1}^{n} \log x_i \\ Ax = b \longrightarrow y \end{cases}$$

Théorème ($\exists$ du chemin central)

Si $\mathcal{F}^s \neq \phi$: $\exists (x, y, s)$: $\begin{cases} A^T y + s = c, & s > 0 \\ Ax = b & x > 0 \end{cases}$

alors 1) (P_μ) a une unique solution x_μ

2) $(\mathcal{D}_\mu)$ a une solution $z_\mu = (x_\mu, y_\mu, s_\mu)$

- unique en (x_μ, s_μ)
- unique en y_μ si A est surjective

Dém 1) existence Soit (x_k) une suite minimisante

Il suffit de montrer que (x_k) est borné

Par l'absurde : sinon, $\exists$ s.s. : $\|x_k\| \to \infty$

- $\dfrac{x_k}{\|x_k\|} \longrightarrow d \neq 0$
- $Ax_k = b \Rightarrow Ad = 0$
- $\underbrace{\dfrac{c^T x_k}{\|x_k\|}}_{\downarrow\ c^T d} - \mu \underbrace{\dfrac{\sum \log (x_k)_i}{\|x_k\|}}_{\downarrow\ 0} \leq \underbrace{\dfrac{C}{\|x_k\|}}_{\downarrow\ 0}$

$\Rightarrow c^T d \leq 0$

Contradiction.

$\mathcal{F}^s_D \neq \phi \Rightarrow \exists (y, s)$: $A^T y + s = c \mid d$ $\quad s > 0$

$0 \leq \underbrace{s^T}_{>0} \underbrace{d}_{\geq 0} \leq 0$ $\quad \underbrace{s^T}_{>0} d = 0 \Rightarrow d = 0$

unicité car $\varphi_\mu(x) = c^T x - \mu \sum_i \log x_i$

est strictement convexe.

$$\nabla \varphi_\mu(x) = c - \mu X^{-1} e \qquad X = \mathrm{Diag}(x_1, \dots, x_n)$$

$$\nabla^2 \varphi_\mu(x) = \mu X^{-2} \succ 0$$

2) car $(\mathcal{O}_\mu) \equiv$ CO de (P_μ)

$$\begin{cases} \min\ c^T x - \mu \sum_i \log x_i \\ Ax = b \quad \leftarrow y \end{cases}$$

$$\Rightarrow \begin{cases} c - \mu X^{-1} e - A^T y = 0 \\ Ax = b \end{cases}$$

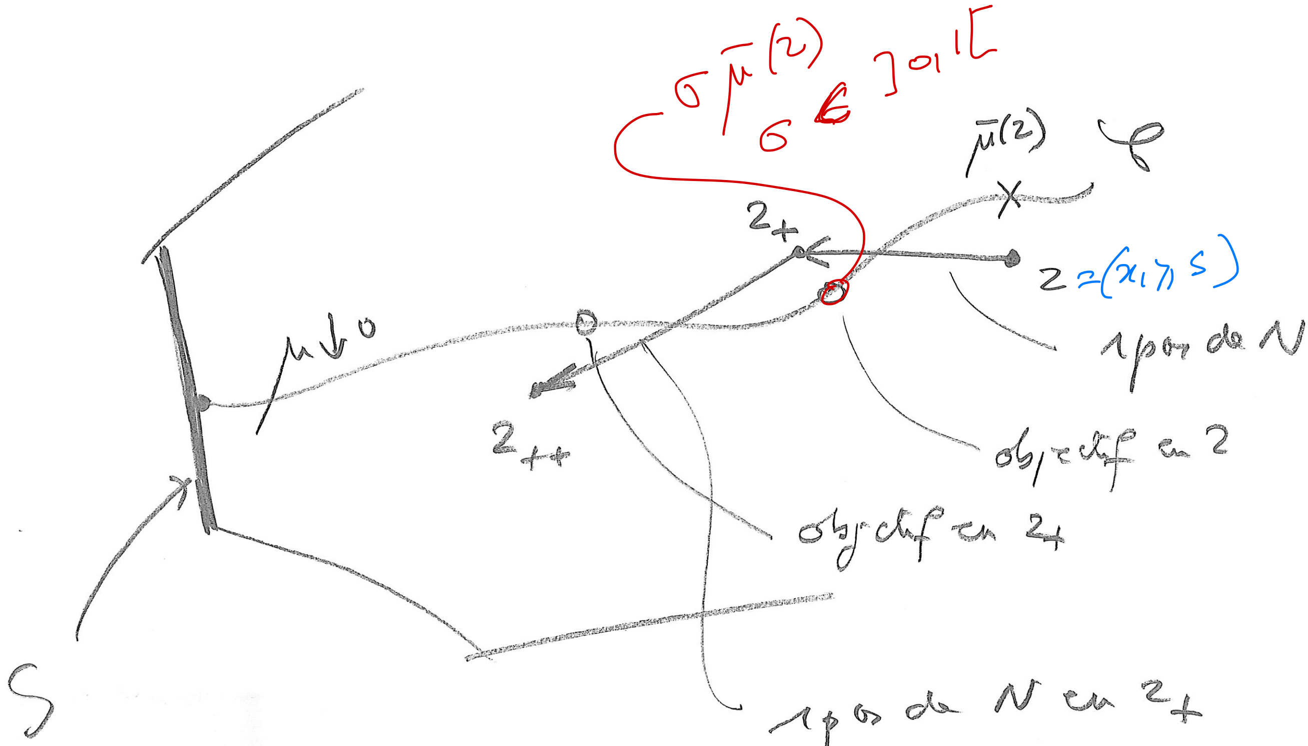
$$s := \mu X^{-1} e$$

$$\Rightarrow \begin{cases} A^T y + s = c \\ Ax = b \\ Xs = \mu e \end{cases} \iff (\mathcal{O}_\mu)$$

$\square$

B Un algo de PI polynomial

Principe : poursuite d'un objectif fuyant



$\Longrightarrow$ on fait des itérations de Newton sur (Θ_μ) avec $\mu \downarrow 0$

Question $\boxed{\text{quel est le point central } z_\mu \text{ le + proche de } z \in \mathcal{F}^s \text{ ?}}$

- mauvaise question car $\mathcal{C}$ n'est pas convexe
- $\Longrightarrow$ on transforme $z \longmapsto \tau(z) = Xs$

  $z_\mu \longmapsto \mu e$ (1 ½-droite)

$$\min_{\mu > 0} \frac{1}{2}\| Xs - \mu e \|_2^2$$

$\rightarrow$ solut $\boxed{\bar\mu(z) = \frac{x^T s}{n}}$

<u>Objectif en $z \in \mathcal{F}^s$</u>

= point central z_μ avec $\mu = \sigma \bar{\mu}(z)$

avec $\sigma \in]0,1[$ mais pas trop petit sinon le pas de Newton sort de $\mathcal{F}^s$

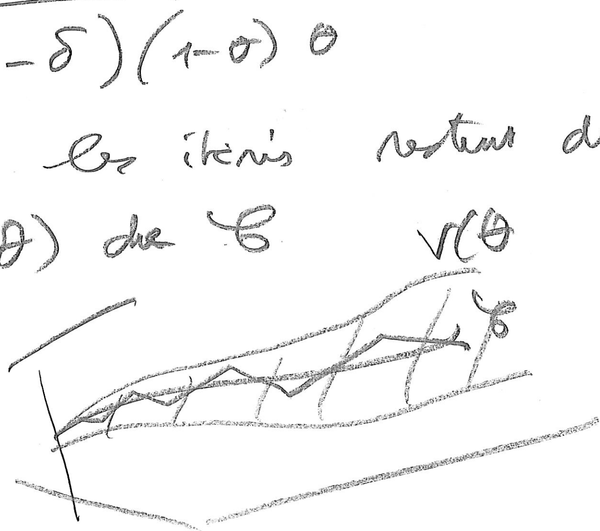
bonne idée $\sigma = 1 - \frac{\delta}{\sqrt{n}}$

directement lié à la complexité de l'algo.
(on ne peut pas faire mieux que $\frac{1}{\sqrt{n}}$; on aimerait prendre $1-\delta$ ⇒ convergence en 1 # d'itérations indép de n !)

$\delta = ?$ Il faut prendre δ et θ tq.

$$\frac{\delta^2 + \theta^2}{(1-\delta)(1-\theta)\theta} \leq \sqrt{8}$$

Alors les itérés restent dans le voisinage $\mathcal{V}(\theta)$ de $\mathcal{C}$



Il reste à voir pourquoi ce σ apporte la polynomialité

— On considère 1 algo "réalisable", "admissible" c-à-d. avec

$$\begin{cases} A^T y + s = c \\ A x = b \end{cases} , \quad \begin{cases} s > 0 \\ x > 0 \end{cases}$$

à chaque itération

linéaire $\Rightarrow$ facile à réaliser

maintenu à chaque itération

— Direction de Newton sur (E_μ)

$$\boxed{\begin{pmatrix} 0 & A^T & I \\ A & 0 & 0 \\ S & 0 & X \end{pmatrix} d = \begin{pmatrix} 0 \\ 0 \\ \sigma \bar{\mu}(z) e - X s \end{pmatrix}}$$

— nouvel itéré

$$\boxed{z_+ = z + d}$$

$z = (x, y, s)$

$$\left\{ \begin{array}{l} A^T y + s = c \\ A x = b \\ X s = \sigma \bar{\mu}(z) e \end{array} \right)$$

l'objectif du pas de Newton

$X = \mathrm{Diag}(x_1, \dots, x_n)$

$S = \mathrm{Diag}(s_1, \dots, s_n)$

Théo (Convergence de l'algo de PI avec petits pas)

1) l'algo converge : $\bar{\mu}(z_k) \longrightarrow 0$ linéairement

2) complexité : on a

$$\bar{\mu}(z_k) \leq \varepsilon\, \bar{\mu}(z_0)$$

dès que $k \geq \dfrac{\sqrt{n}}{\delta} \log \dfrac{1}{\varepsilon}$

Dém 1) $z := z_k$

$$\bar{\mu}(\underbrace{z+d}_{\to z_{k+1}}) = \frac{1}{n}(x+dx)^T(s+ds)$$

$$= \bar{\mu}(z) + \frac{1}{n}\left(s^T dx + x^T ds\right) + \frac{1}{n}\underbrace{dx^T ds}_{=\,0}$$

$e = \begin{pmatrix} 1 \\ \vdots \\ 1 \end{pmatrix}$

$e^T \,|\, S\,dx + X\,ds = \sigma \bar{\mu}(z) e - Xs$ $\qquad dx^T A^T dy = 0$

$$\frac{s^T dx + x^T ds}{n} = \frac{n\sigma\bar{\mu}(z) - x^T s}{n}$$

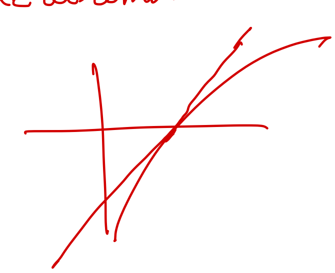
$$= \bar{\mu}(z) + \sigma\bar{\mu}(z) - \bar{\mu}(z)$$

$$= \sigma\bar{\mu}(z)$$

$$= \left(1 - \frac{\delta}{\sqrt{n}}\right)\bar{\mu}(z) \longrightarrow 0 \text{ linéairement}$$

2) $\bar{\mu}_k = \left(1 - \dfrac{\delta}{\sqrt{n}}\right)\bar{\mu}_{k-1}$

$$\log \bar{\mu}_k = \log\left(1 - \frac{\delta}{\sqrt{n}}\right) + \log \bar{\mu}_{k-1}$$



$$\log(1+t) \leq \underbrace{\log 1}_{0} + t\,\underbrace{\log'(1)}_{1}$$

$\leq t$

$$\leq -\frac{\delta}{\sqrt{n}} + \log \bar{\mu}_{k-1}$$

$$\leq -\frac{\delta k}{\sqrt{n}} + \log \bar{\mu}_0$$

$$\log \frac{\bar{\mu}_k}{\bar{\mu}_0} \leq -\frac{k\delta}{\sqrt{n}} \leq \log \varepsilon \quad \text{si}$$

$$k \geq \frac{\sqrt{n}}{\delta} \log \varepsilon^{-1}$$

$$\Rightarrow \text{on a} \quad \frac{\bar{\mu}_k}{\bar{\mu}_0} \leq \varepsilon \quad \text{si}$$

$\square$