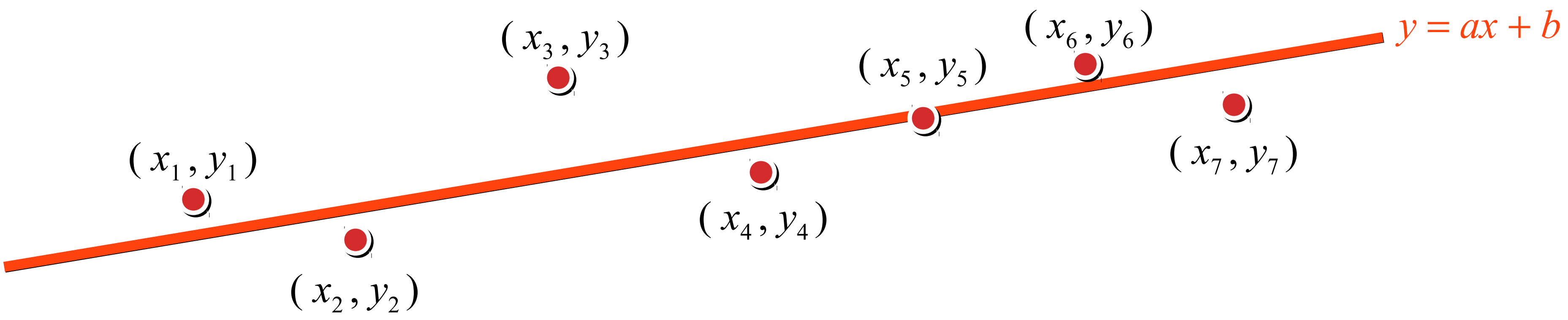
Méthodes numériques approximation et optimisation

INSTITUT d'OPTIQUE
GRADUATE SCHOOL
ParisTech

Objectifs

- Données mesurées vers modèle numérique
  - Étudier le comportement
  - Utiliser dans une simulation
  - Supprimer le bruit
  - …
- Meilleure approximation



Que veut dire meilleure ?

- Ne pas trop s'éloigner
  - Notion de distance
  - Notion de norme
- Définition de **norme**

$$\boldsymbol{x} = \begin{pmatrix} x_1 & \dots & x_N \end{pmatrix}^{\mathrm{T}}$$

  - Séparation

$$\|\boldsymbol{x}\| = 0 \Leftrightarrow \forall\, i = 1..\mathrm{N}\,,\, x_i = 0$$

  - Homogénéité

$$\|\lambda\, \boldsymbol{x}\| = |\lambda|\, \|\boldsymbol{x}\|$$

  - Inégalité triangulaire

$$\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$$

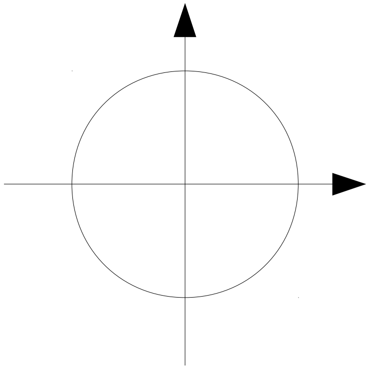
Rappel : norme euclidienne

$$\left\|(x_1, \ldots, x_N)\right\|_2 = \left(\sum_{i=1}^{N} x_i^2\right)^{\frac{1}{2}}$$

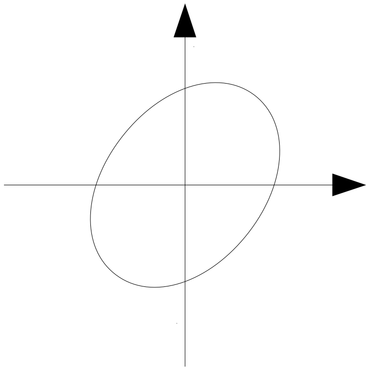
- A partir d'un produit scalaire

$$\langle \boldsymbol{x}, \boldsymbol{y} \rangle = \sum_{i=1}^{N} x_i y_i = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{y}$$

$$\|\boldsymbol{x}\|_2^2 = \langle \boldsymbol{x}, \boldsymbol{x} \rangle = \boldsymbol{x}^{\mathrm{T}} \boldsymbol{x}$$



- Erreur ≈ distance moyenne
- Produit scalaire généralisée
  - Q matrice symétrique définie positive

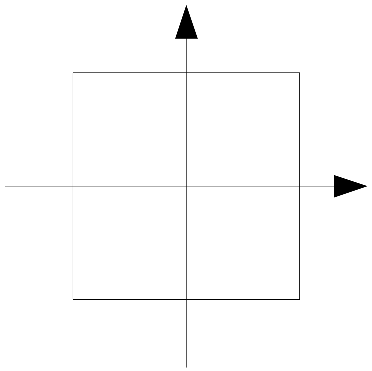


$$\|\boldsymbol{x}\|_{\mathrm{Q}}^2 = \langle \boldsymbol{x}, \boldsymbol{x} \rangle_{\mathrm{Q}} = \boldsymbol{x}^{\mathrm{T}} \mathrm{Q}\, \boldsymbol{x}$$

Rappel : norme infinie

- Maximum des valeurs absolues

$$\left\|\left(x_{1,}\ldots,x_N\right)\right\|_\infty = \max\left\{\left|x\right|_{i=1..\mathrm{N}}\right\}$$

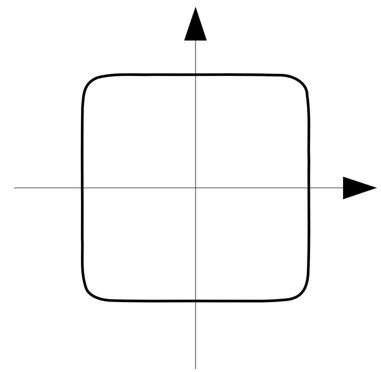


- Sur les erreurs = borne maximale

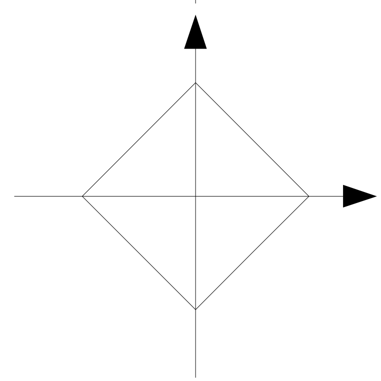
Définition : norme p

- Généralisation de la norme euclidienne

$$\left\|\left(x_{1,}\dots,x_N\right)\right\|_p=\left(\sum_{i=1}^{N}\left|x_i\right|^p\right)^{\frac{1}{p}}$$



$\|\,\|_{10}$



$\|\,\|_{1}$

Méthodes linéaires

- Trouver la solution optimum se réduit à
  - Système linéaire

$$\mathrm{A}\,\boldsymbol{x}-\boldsymbol{b}=0$$

  - Des contraintes linéaires

$$\boldsymbol{c}_i^{\mathrm{T}}\boldsymbol{x}-\boldsymbol{d}_i\leq 0 \qquad\qquad \boldsymbol{c}_j^{\mathrm{T}}\boldsymbol{x}-\boldsymbol{d}_j=0$$

- Résolution simplifiée
  - En général : calcul matriciel
  - En général : solution unique

Moindres carrés

- Mesures $(\boldsymbol{x}_{m=1..M}, \boldsymbol{y}_{m=1..M})$
- Fonction d'approximation
  - Paramètres $\boldsymbol{v}$ $\qquad \boldsymbol{v} = \begin{pmatrix} v_1 & \dots & v_K \end{pmatrix}^{\mathrm{T}}$
  - Paramètres linéaires $\qquad \boldsymbol{y} = \mathbf{f}_{\boldsymbol{v}}(\boldsymbol{x}) = \sum_{k=1}^{K} v_k \mathbf{f}_k(\boldsymbol{x})$
  - Exemples 2D: droite $y = ax + b$

$$y = \mathrm{f}_{(a,b)}(x) = ax + b$$

Moindres carrés

- Erreur minimum pour la norme euclidienne

$$\min_{\boldsymbol{v}} \sum_{m=1}^{M} \left\| \boldsymbol{y}_m - \mathbf{f}_{\boldsymbol{v}}\left(\boldsymbol{x}_m\right) \right\|_2^2$$

- Existence d'une solution unique :
  - $M \geq K$
  - Éviter des problèmes de conditionnement
    - Ne contient pas la solution triviale $\boldsymbol{v}=\mathbf{0}$
    - Exemple de problème de conditionnement : droite implicite

$$0 = \mathrm{f}_{(a,b,c)}(x,y) = ay + bx + c$$

Moindres carrés : résolution

- Fonction à minimiser :
  - Toujours positive (pas de borne sur les paramètres)
  - Quadratique (norme euclidienne)
- Minimum atteint là où la dérivé s'annule

$$\forall k = 1..\mathrm{K}\,,\ \frac{\partial}{\partial v_k} \sum_{m=1}^{M} \left\| \boldsymbol{y}_m - \mathbf{f}_{\boldsymbol{v}}\left(\boldsymbol{x}_m\right) \right\|_2^2 = 0$$

- Système linéaire A matrice symétrique

$$\mathrm{A}\,\boldsymbol{v} = \boldsymbol{b}$$

Système

y et f_v sont des scalaires

$$b_k = \sum_{m=1}^{M} \mathrm{f}_k(\boldsymbol{x}_m)\, y_m$$

$$a_{ki} = \sum_{m=1}^{M} \mathrm{f}_k(\boldsymbol{x}_m)\,\mathrm{f}_i(\boldsymbol{x}_m)$$

$\boldsymbol{y}$ et $\mathbf{f}_v$ sont des vecteurs

$$b_k = \sum_{m=1}^{M} \left\langle \mathbf{f}_k(\boldsymbol{x}_m), \boldsymbol{y}_m \right\rangle$$

$$a_{ki} = \sum_{m=1}^{M} \left\langle \mathbf{f}_k(\boldsymbol{x}_m), \mathbf{f}_i(\boldsymbol{x}_m) \right\rangle$$

$$\boldsymbol{b} = \mathrm{F}\,\boldsymbol{y}$$

$$\mathrm{A} = \mathrm{F}\,\mathrm{F}^{\mathrm{T}}$$

Rajouter des conditions

- Pour régulariser
  - Améliorer le conditionnement
  - Exemple de la droite implicite

$$\min_{\boldsymbol{v}} \sum_{m=1}^{M} \left\| \mathrm{f}_{\boldsymbol{v}}\left(\boldsymbol{x}_m\right) \right\|_2^2 + \epsilon \left(\sum_{k=1}^{K} v_k - 1 \right)^2$$

- D'autres contraintes
  - Continuité, lissage
  - Domaine de définition

Multiplicateur de Lagrange

- Approche générique
  - Un fonction à minimiser $\min_{\boldsymbol{v}} \mathrm{f}(\boldsymbol{v})$
  - Une contrainte $\mathrm{g}(\boldsymbol{v}) = b$
  - Nouvel objectif $\min_{\boldsymbol{v},\lambda} \mathrm{f}(\boldsymbol{v}) + \lambda\left(\mathrm{g}(\boldsymbol{v}) - b\right)$

- Minimum atteint lorsque

$$\frac{\partial}{\partial v_k}\left(\mathrm{f}(\boldsymbol{v}) + \lambda\,\mathrm{g}(\boldsymbol{v})\right) = 0$$

$$\frac{\partial}{\partial \lambda} \mathrm{f}(\boldsymbol{v}) + \lambda\left(\mathrm{g}(\boldsymbol{v}) - b\right) = 0 = \mathrm{g}(\boldsymbol{v}) - b$$

Cas des moindres carrés

- Contrainte linéaire

$$\boldsymbol{c}_j^{\mathrm{T}} \boldsymbol{v} = b_j$$

- Nouvelle fonction d'optimisation

$$\min_{\boldsymbol{v}} \sum_{m=1}^{M} \left\| \mathbf{f}_{\boldsymbol{v}}\left(\boldsymbol{x}_m\right) \right\|_2^2 + \sum_{j=1}^{J} \lambda_j \left(\boldsymbol{x}_j^{\mathrm{T}} \boldsymbol{v} - b_j \right)$$

- Solution d'un nouveau système linéaire

Moindres carrés

- Norme euclidienne
  - En moyenne meilleur
  - Robuste au bruit
  - Possibilité d'adapter la norme
  - Se réduit à une inversion de matrice

- Mais, quelle est l'erreur maximal
  - Norme infinie

- D’autres contraintes (inégalités)

Programmation linéaire

- Minimiser une norme infinie

$$\min_{\boldsymbol{v}} \max_{m=1}^{M} \left\| y_m - \mathrm{f}_{\boldsymbol{v}}\left(\boldsymbol{x}_m\right) \right\|_{\infty}$$

$$\min_{\boldsymbol{v}} \left\| \boldsymbol{Y} - \mathbf{f}_{\boldsymbol{v}}\left(\boldsymbol{X}\right) \right\|_{\infty}$$

$$\Leftrightarrow \begin{array}{l} \min_{\boldsymbol{v},\,\epsilon} \epsilon \\ \text{subject to} \begin{cases} \epsilon \geq 0 & \\ -y_m + \mathrm{f}_{\boldsymbol{v}}\left(\boldsymbol{x}_m\right) + \epsilon \geq 0 & \forall\, m \\ y_m - \mathrm{f}_{\boldsymbol{v}}\left(\boldsymbol{x}_m\right) - \epsilon \geq 0 & \forall\, m \end{cases} \end{array}$$

Programmation linéaire

- Programme linéaire
  - Objectif : un produit scalaire
  - Contraintes : égalités, inégalité linéaires

$$\min_{v} \boldsymbol{d}^{\mathrm{T}} \boldsymbol{v}$$
$$\text{subject to} \quad \boldsymbol{c}_m^{\mathrm{T}} \boldsymbol{v} \leq b_m$$

- Résolution
  - Nombre d'itérations ~ O(nombre de contraintes)

Simplexe : forme standard

$$\max_{\boldsymbol{v}} \boldsymbol{d}^{\mathrm{T}} \boldsymbol{v}$$

$$\text{subject to} \begin{cases} \boldsymbol{c}_m^{\mathrm{T}} \boldsymbol{v} + e_m = b_m & \forall\, m \\ v_k \geq 0 & \forall\, k \\ e_m \geq 0 & \forall\, m \end{cases}$$

$$\text{with} \begin{cases} \boldsymbol{v} = \begin{pmatrix} v_1 & .. & v_k \end{pmatrix}^{\mathrm{T}} \\ \boldsymbol{d} = \begin{pmatrix} d_1 & .. & d_k \end{pmatrix}^{\mathrm{T}} \\ \boldsymbol{c}_m = \begin{pmatrix} c_{m1} & .. & c_{mk} \end{pmatrix}^{\mathrm{T}} \end{cases}$$

Simplexe : vers la forme standard

- Si l'objectif est de minimiser $\boldsymbol{d}^{\mathrm{T}}\boldsymbol{v}$
  - maximiser $-\boldsymbol{d}^{\mathrm{T}}\boldsymbol{v}$
- Si on a des inégalités
  - $\boldsymbol{d}^{\mathrm{T}}\boldsymbol{v} \leq b$: transformer en $\boldsymbol{d}^{\mathrm{T}}\boldsymbol{v} + s = b$ et $s \geq 0$
  - $\boldsymbol{d}^{\mathrm{T}}\boldsymbol{v} \geq b$: transformer en $\boldsymbol{d}^{\mathrm{T}}\boldsymbol{v} - s = b$ et $s \geq 0$
- Si pas de contrainte sur v_k
  - Remplacer v_k par $y_k - z_k$
  - Ajouter $y_k \geq 0$ et $z_k \geq 0$

Processus itératif : exemple

- Maximiser $7x_1 + 5x_2$

  Sachant que :

$$x_1 \leq 300$$

$$x_2 \leq 400$$

$$x_1 + x_2 \leq 500$$

$$2x_1 + x_2 \leq 700$$

$$x_1 \geq 0$$

$$x_2 \geq 0$$

Mise à la forme standard

- Maximiser $7x_1 + 5x_2$

  Sachant que

  $x_1 + e_1 = 300$

  $x_2 + e_2 = 400$

  $x_1 + x_2 + e_3 = 500$

  $2x_1 + x_2 + e_4 = 700$

  $x1 \geq 0\ ;\ \ x2 \geq 0$

  $e_i \geq 0$

Tableau de départ du simplexe

T1	x_1	x_2	e_1	e_2	e_3	e_4	b
e_1	1	0	1	0	0	0	300
e_2	0	1	0	1	0	0	400
e_3	1	1	0	0	1	0	500
e_4	2	1	0	0	0	1	700
	7	**5**	**0**	**0**	**0**	**0**	**0**

Critère d'arrêt : coefficients de la dernière ligne négatifs

Tableau de départ du simplexe

T1	x_1	x_2	e_1	e_2	e_3	e_4	b
e_1	1	0	1	0	0	0	300
e_2	0	1	0	1	0	0	400
e_3	1	1	0	0	1	0	500
e_4	2	1	0	0	0	1	700
	7	**5**					**0**

Tableau de départ du simplexe

T1	x_1	x_2	e_1	e_2	e_3	e_4	b
x_1	1	0	1	0	0	0	300
e_2	0	1	0	1	0	0	400
e_3	0	1	-1	0	1	0	200
e_4	0	1	-2	0	0	1	100
	0	**5**	**-7**	**0**	**0**	**0**	**-2100**

Tableau de départ du simplexe

T1	x_1	x_2	e_1	e_2	e_3	e_4	b
x_1	1	0	1	0	0	0	300
e_2	0	1	0	1	0	0	400
e_3	0	1	-1	0	1	0	200
e_4	0	1	-2	0	0	1	100
	0	**5**	**-7**	**0**	**0**	**0**	**-2100**

Tableau de départ du simplexe

T1	x_1	x_2	e_1	e_2	e_3	e_4	b
x_1	1	0	1	0	0	0	300
e_2	0	0	2	1	0	0	300
e_3	0	0	1	0	1	0	100
x_2	0	1	-2	0	0	1	100
	0	**0**	**3**	**0**	**0**	**-5**	**-2600**

Tableau de départ du simplexe

T1	x_1	x_2	e_1	e_2	e_3	e_4	b
x_1	1	0	1	0	0	0	300
e_2	0	0	2	1	0	0	300
e_3	0	0	1	0	1	0	100
x_2	0	1	-2	0	0	1	100
	0	**0**	**3**	**0**	**0**	**-5**	**-2600**

Tableau de départ du simplexe

T1	x_1	x_2	e_1	e_2	e_3	e_4	b
x_1	1	0	0	0	-1	0	200
e_2	0	0	0	1	-2	0	100
e_1	0	0	1	0	1	0	100
x_2	0	1	0	0	2	1	300
	0	**0**	**0**	**0**	**-3**	**-5**	**-2900**

Solution optimale

En base :

$\boldsymbol{x_1 = 200}$

$e_2 = 100$

$e_1 = 100$

$\boldsymbol{x_2 = 300}$

$e_3 = e_4 = 0$ (hors base)

Max $\boldsymbol{= 2900}$

Interprétation géométrique

MAXIMISER $3\,x_1 + 5\,x_2$

Contraintes :

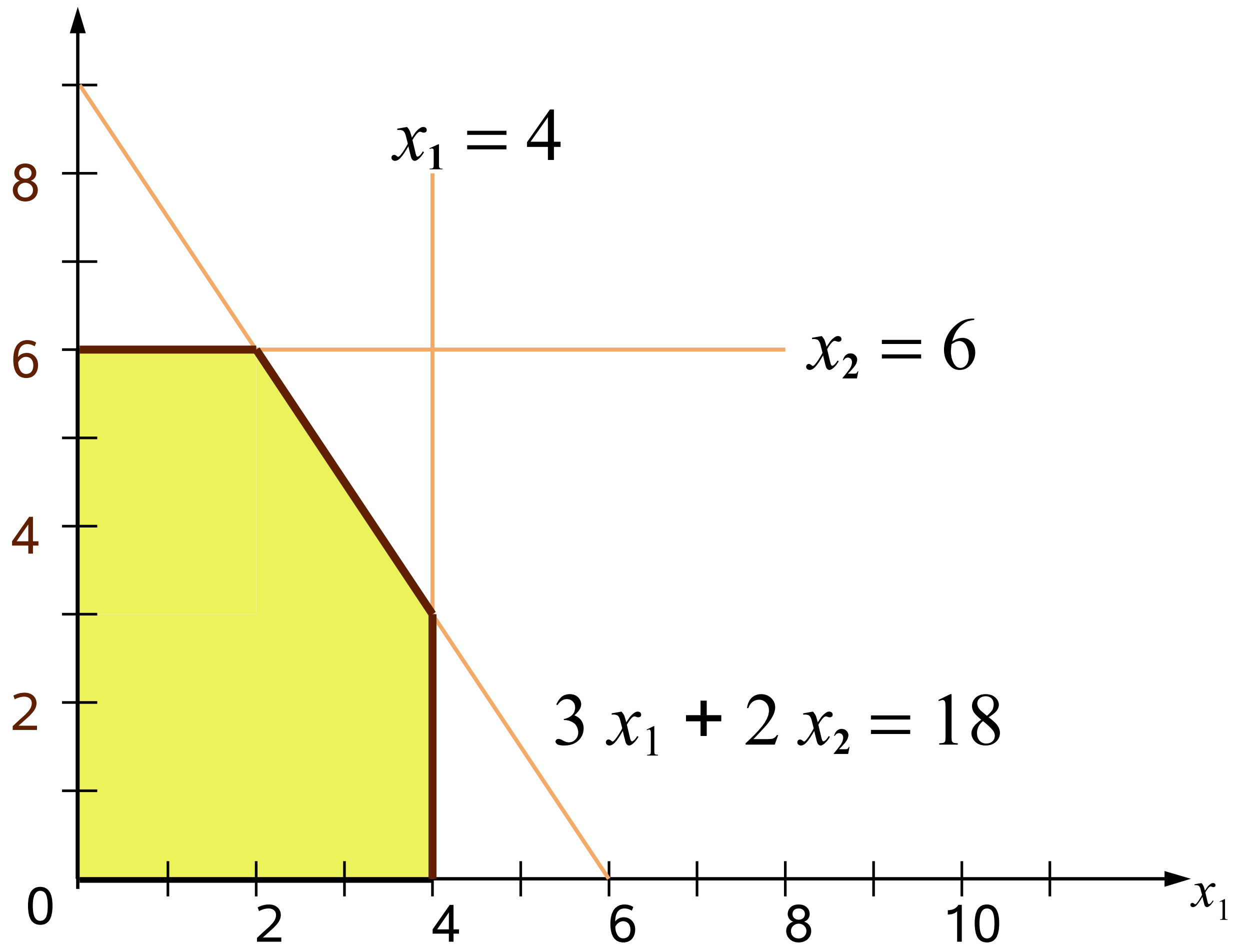
$x_1 \leq 4$

$x_2 \leq 6$

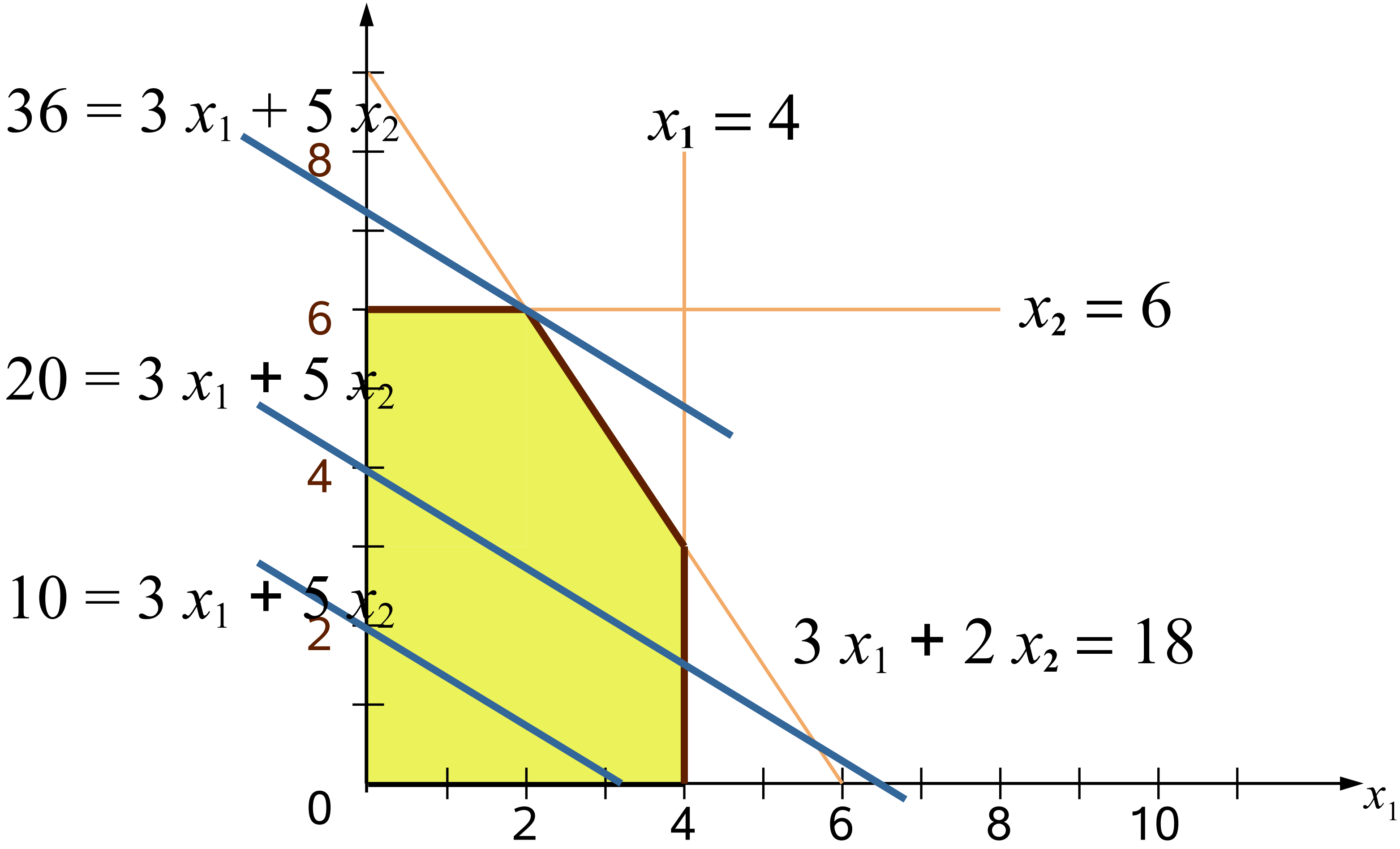
$3\,x_1 + 2\,x_2 \leq 18$

$x_1 \geq 0\ ;\ x_2 \geq 0$

Interprétation géométrique : domaine



Interprétation géométrique : objectif



Programmation quadratique

- Terme à minimiser = forme quadratique

$$\min_v \boldsymbol{v}^{\mathrm{T}} \mathrm{Q}^{\mathrm{T}} \boldsymbol{v} + \boldsymbol{d}^{\mathrm{T}} \boldsymbol{v}$$

$$\text{subject to } \begin{cases} \boldsymbol{c}_j^{\mathrm{T}} \boldsymbol{v} = b_j \\ \boldsymbol{c}_m^{\mathrm{T}} \boldsymbol{v} \leq b_m \end{cases}$$

- Principe de résolution
  - Solution aux moindres carrés avec les égalités
  - Si les inégalités ne sont pas toutes vérifiées
    - En transformer une en égalité et recommencer

Moindres carrés

- Minimiser une norme euclidienne

$$\min_V \sum_{m=1}^{M} \left\| \boldsymbol{y}_m - \mathbf{g}_V\left(\boldsymbol{x}_m\right) \right\|_2^2$$

$$\min_V \left\| \boldsymbol{Y} - \mathbf{g}_V\left(\boldsymbol{X}\right) \right\|_2^2$$

$$\boldsymbol{Y} = \begin{pmatrix} \boldsymbol{y}_1 & \ldots & \boldsymbol{y}_M \end{pmatrix}^{\mathrm{T}}$$

$$\mathbf{g}_V\left(\boldsymbol{X}\right) = \begin{pmatrix} \mathbf{g}_V\left(\boldsymbol{x}_1\right) & \ldots & \mathbf{g}_V\left(\boldsymbol{x}_M\right) \end{pmatrix}^{\mathrm{T}}$$

INSTITUT d'OPTIQUE GRADUATE SCHOOL ParisTech