MARO 018 - Optimization & Operational Research

Arnaud Vandaele

Faculté Polytechnique
Service MAthématique et Recherche Opérationnelle

UMONS

Année académique 2024-2025

Séance 3 :
Optimisation continue et combinatoire : liens et complexité.

Séance 3 – Optimisation continue et combinatoire : liens et complexité.

Tables des matières

Plan

Optimisation quadratique : les différents cas

$$\boxed{\min_x \frac{1}{2}x^T Qx - c^T x + p \quad \text{s.t.} \quad Ax \leq b \;\text{ et }\; Mx = d.}$$

Les différents cas :

1. Sans contraintes : $A = 0,\ b = 0,\ M = 0,\ d = 0$
2. Avec des contraintes d'égalités : $A = 0,\ b = 0$
3. Avec des contraintes d'inégalités : $M = 0,\ d = 0$

Optimisation quadratique : sans contraintes

Avec $Q \in \mathbb{R}^{n \times n}$ symétrique et $c \in \mathbb{R}^n$, le problème est :

$$\boxed{\min_x \frac{1}{2} x^T Q x - c^T x + p.}$$

Ce problème peut être analysé par des méthodes classiques d'algèbre linéaire :

1. **Stationnarité.** Le gradient de f est $\nabla f(x) = Qx - c$, ce qui signifie que chercher un point stationnaire revient à résoudre
$$Qx = c.$$
   - Dans le cas idéal, Q est inversible.
   - Q peut également être de rang $< n$.
   - Il peut ne pas exister de point stationnaire.
2. Caractère du point stationnaire (s'il en existe un)
   - Si $Q \succeq 0$, alors il existe un minimum global
   - Sinon, on a soit un maximum, soit une matrice Q indéfinie, et, dans les deux cas, le problème ne possède pas de minimum.

Dans les deux points, le calcul des valeurs propres de Q est nécessaire.
Ceci peut être fait en $\mathcal{O}(n^3)$.

Optimisation quadratique : avec contraintes d'égalités

Avec $Q \in \mathbb{R}^{n \times n}$ symétrique, $c \in \mathbb{R}^n$ et $M \in \mathbb{R}^{m \times n}$, le problème est :

$$\boxed{\min_x \frac{1}{2} x^T Q x - c^T x + p \quad \text{s.t.} \quad Mx = d.}$$

Les conditions KKT mènent à :

$$\begin{pmatrix} Q & -M^T \\ M & 0 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} c \\ d \end{pmatrix}.$$

La résolution de ce système peut être effectuée en $\mathcal{O}((m+n)^3)$.

Optimisation quadratique : avec contraintes d'inégalités

$$\boxed{\min_x \frac{1}{2} x^T Q x - c^T x + p \quad \text{s.t.} \quad Ax \leq b.}$$

Les conditions KKT mènent à :

$$Qx - c + A^T y = 0 \quad \text{et} \quad y^T(b - Ax) = 0 \quad \text{et} \quad y \geq 0.$$

Lorsque $Q \succeq 0$, le problème est convexe et s'il existe un point de slater, alors les KKT deviennent nécessaires et suffisantes.

Idée d'algorithme : tester les 2^m partitionnements $\mathcal{K} \cup \bar{\mathcal{K}} = \{1, ..., m\}$ possibles :

- Active set : $\mathcal{K} = \{j \in \{1, ..., m\} \ : \ A(j, :)x = b_j\}$.
- Passive set : $\bar{\mathcal{K}} = \{j \in \{1, ..., m\} \ : \ A(j, :)x < b_j\}$.

Mais la complexité est exponentielle en m.

Autre idée d'algorithme : méthodes de points intérieurs

Optimisation quadratique (QP) : difficile en général

$$\boxed{\min_x \frac{1}{2}x^T Qx - c^T x + p \quad \text{s.t.} \quad Ax \leq b.}$$

- QP peut être résolu en temps polynomial pour $Q \succeq 0$
- QP est NP-hard [Sahni, 1974]
- QP est dans NP [Vavasis, 1990]
- QP avec une seule valeur propre négative est NP-hard [Pardalos, Vavasis, 1990]

(il peut y avoir de nombreux minima locaux pour ces problèmes)

Qu'en est-il de $\text{QP}_{\geq 0}$?

$$\boxed{\min_x \frac{1}{2}x^T Qx - c^T x + p \quad \text{s.t.} \quad x \geq 0.}$$

Réduction à partir du problème Partition

Partition est un problème NP-hard :

INSTANCE: Un ensemble S d'entiers positifs s_i.

QUESTION: trouver une partition $S = S_1 \cup S_2$, $S_1 \cap S_2 = \emptyset$ telle que $\sum_{i \in S_1} s_i = \sum_{i \in S_2} s_i$.

Example: $S = \{3, 1, 2\}$.

Que peut-on dire de la complexité de $\mathrm{QP}_{\geq 0}$?

INSTANCE: $Q \in \mathbb{R}^{n \times n}$, $c \in \mathbb{R}^n$ et $p \in \mathbb{R}$.

QUESTION: trouver $x \geq 0$ tel que $\frac{1}{2} x^T Q x - c^T x + p \leq 0$.

- A partir de toute instance S de Partition,

  on décrit la création d'une instance de $\mathrm{QP}_{\geq 0}$ en temps polynomial.
  - $n = 2|S|$ (avec $x = \begin{pmatrix} x_1^+ & x_1^- & \dots & x_n^+ & x_n^- \end{pmatrix}^T$)
  - Q, c et p représentent la fonction quadratique suivante

$$f(x) = \sum_{i \in S} \left(\left(x_i^+ + x_i^- - 1 \right)^2 + x_i^+ x_i^- \right) + \left(\sum_{i \in S} x_i^+ s_i - \frac{1}{2} \sum_{i \in S} s_i \right)^2 .$$

- Il faut prouver: (une solution $\Rightarrow$ une solution) et (une solution $\Leftarrow$ une solution)

Reduction - Preuve

Instance de Partition: $S = \{3, 1, 2\}$.

Instance de $\text{QP}_{\geq 0}$:

$$Q = \begin{pmatrix} 11 & 3 & 6 & 0 & 12 & 0 \\ 3 & 2 & 0 & 0 & 0 & 0 \\ 6 & 0 & 3 & 3 & 4 & 0 \\ 0 & 0 & 3 & 2 & 0 & 0 \\ 12 & 0 & 4 & 0 & 6 & 3 \\ 0 & 0 & 0 & 0 & 3 & 2 \end{pmatrix} \quad , \quad c = \begin{pmatrix} -20 \\ -2 \\ -8 \\ -2 \\ -14 \\ -2 \end{pmatrix} \quad \text{et } p = 12.$$

- Preuve de (une solution $x \Rightarrow$ une solution (S_1, S_2)).
  Par construction, $f(x) \geq 0$. Comme $f(x) \leq 0$, on $f(x) = 0$, donc:
  $\sum_{i \in S} \left(\left(x_i^+ + x_i^- - 1 \right)^2 + x_i^+ x_i^- \right) = 0 \quad \Rightarrow \quad \{x_i^+, x_i^-\} \in \{0, 1\}$.
  $\sum_{i \in S} x_i^+ s_i - \frac{1}{2} \sum_{i \in S} s_i = 0 \quad \Rightarrow \quad$ on a trouvé (S_1, S_2).
- Preuve de (une solution $(S_1, S_2) \Rightarrow$ une solution x).
  Pour $i \in S_1$, on assigne $x_i^+ = 1$ et $x_i^- = 0$.
  Pour $i \in S_2$, on assigne $x_i^+ = 0$ et $x_i^- = 1$.
  L'objectif quadratique vaut alors 0. $\quad \Rightarrow \quad$ on a trouvé x.
- $\text{QP}_{\geq 0}$ est NP-hard.

Plan

Optimisation quadratique binaire : linéarisation

$$\boxed{\min_x \frac{1}{2} x^T Q x - c^T x + p \quad \text{avec} \quad x \in \{0,1\}^n.}$$

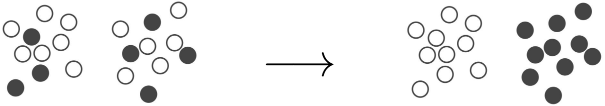
Rappel : $x^T Q x = \sum_{i=1}^n \sum_{j=1}^n Q_{ij} x_i x_j$

Ce problème possède des termes non-linéaires $x_i x_j$.
Cependant, il est possible de le formuler comme un problème linéaire en variables binaires à l'aide de : une variable binaire et 3 contraintes par terme quadratique :

$$\text{pour tout terme } xy \quad \rightarrow \quad z = xy \quad \rightarrow \quad x \geq z,\ y \geq z \text{ et } x + y - 1 \leq z.$$

Problème : Cluster Analysis

- Soit un ensemble de N points de $\mathbb{R}^2$, décrits par une matrice $D \in \mathbb{R}^{N \times 2}$.
- On souhaite subdiviser ces points en deux clusters de façon à ce que les points dans le même cluster soient le plus similaires les uns aux autres.
- Cette similarité sera ici considérée via la distance euclidienne entre les points.



Modélisation :

- Pour chacun des points $i = 0, ..., N-1$, on a une variable binaire qui indique si le point i appartient au premier ($x_i = 0$) ou second cluster ($x_i = 1$).
- On peut remarquer que si les point i et j appartiennent au même cluster, la quantité suivante vant 1 :

$$x_i x_j + (1 - x_i)(1 - x_j).$$

  Pareillement, si i et j ne sont pas dans le même cluster, la quantité suivante vaut 1 :

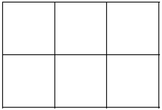
$$x_i(1 - x_j) + (1 - x_i)x_j.$$

- En définissant d_{ij} comme la distance euclidienne entre les points i et j, la minimisation de la fct. obj. suivante aura tendance à placer les points éloignés dans des clusters différents :

$$f(x) = \sum_{i=0}^{N-1} \sum_{j=i+1}^{N-1} d_{ij}(x_i x_j + (1 - x_i)(1 - x_j)) - d_{ij}(x_i(1 - x_j) + (1 - x_i)x_j).$$

Problème : Implantation d'ateliers

- Dans un atelier, on souhaite placer n machines (qui possède le même encombrement) sur n emplacements possibles.
- On connaît la distance d_{kl} entre deux emplacements k et l, et on connaît le flux de produits f_{ij} entre la machine i et la machine j.



Modélisation :

- On définit les variables binaires $x_i k$ qui déterminent si la machine i est localisée à l'emplacement k.
- On cherche à minimiser le coût de déplacement des produits entre les machines :

$$\sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k=1}^{n} \sum_{l=1}^{n} f_{ij} d_{kl} x_{ik} x_{jl}$$

Plan

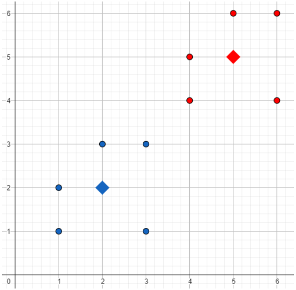
K-means est un problème d'optimisation !

INPUT: $X \in \mathbb{R}^{m \times n}$, représentant n points de $\mathbb{R}^m$ et $k \leq n$.

QUESTION: trouver une partition $C_1, ..., C_k$ des points tel que

$$\min_{C_1,\dots,C_k} \sum_{t=1}^{k} \|X(:, C_t) - W(:, t)\boldsymbol{e}^T\|_F^2,$$

minimisant la somme des distances au carré entre les points $X(:, j)$ et le centroïde $W(:, t) = \frac{1}{|C_t|} \sum_{i \in C_t} X(:, i)$ de son cluster.



INPUT: $k = 2$ et

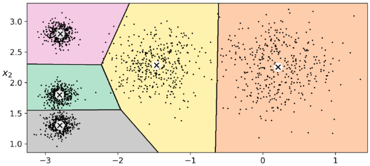
$$X = \begin{pmatrix} 1 & 1 & 2 & 3 & 3 & 4 & 4 & 5 & 6 & 6 \\ 1 & 2 & 3 & 1 & 3 & 4 & 5 & 6 & 4 & 6 \end{pmatrix}$$

SOLUTION: $C_1 = \{1, 2, 3, 4\}$, $C_2 = \{5, 6, 7, 8\}$

$$We^T = \begin{pmatrix} 2 & 2 & 2 & 2 & 2 & 5 & 5 & 5 & 5 & 5 \\ 2 & 2 & 2 & 2 & 2 & 5 & 5 & 5 & 5 & 5 \end{pmatrix}$$

Un algorithme **EXACT** en $\mathcal{O}(n^{mk+1})$

Applications of weighted Voronoi diagrams and randomization to variance-based k-clustering. Inaba, Katoh, Imai, 1994.



- Un k-clustering optimal est une partition de Voronoi
- Les k-partitions de Voronoi de n points peuvent être énumérées en $\mathcal{O}(n^{mk+1})$.

Qu'est-ce que cela implique ?

- Si la dimension m et le nombre de clusters k sont fixés (ils ne font pas partie de l'input), alors k-means peut être résolu en temps polynomial !

Mais ...

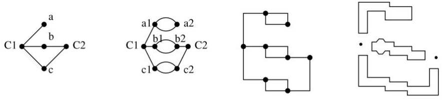
- Qu'en est-il quand uniquement la dimension m est fixée?
- Qu'en est-il quand uniquement le nombre de clusters k est fixé?

Quand m est fixé, mais pas k

Mauvaise nouvelle: NP-hard, même quand $m = 2$.
(The Planar k-Means Problem is NP-Hard. Mahajan, Nimbhorkar, Varadarajan (2009))
(The hardness of kmeans clustering in the plane. Vattani (2009))

Une première réduction : PLANAR 3-SAT (problème NP-hard):

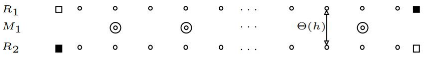
INPUT:	A 3-CNF formula made of n variables and m clauses whose incidence graph is planar.
QUESTION:	determine if the formula is satisfiable.



The graph G — The graph H — The embedding of H — The final layout

Une seconde réduction : EXACT COVER BY 3-SETS (problème NP-hard):

INSTANCE:	Un set X t.q. $|X| = 3q$ et une collection C de subsets de 3-elements de X.
QUESTION:	trouver des subsets dans C tels que leur union donne X.

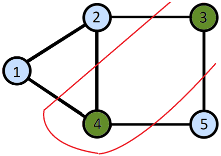


Quand k est fixé, mais pas m

Mauvaise nouvelle: NP-hard, même quand $k = 2$. *(NP-hardness of Euclidean SOS clustering (2009).)*

Réduction de : DENSEST-CUT (problème NP-hard) :

INPUT:	Un graphe $G = (E, V)$.
QUESTION:	trouver une partition $(V, \bar{V})$ des sommets maximisant $\frac{\lvert E(V,\bar{V})\rvert}{\lvert V\rvert . \lvert \bar{V}\rvert}$.



$\rightarrow$

$n = |V|$,
$m = |E|$,
$k = 2$

$$X = \begin{array}{c} \\ (1,2) \\ (2,3) \\ (1,4) \\ (2,4) \\ (3,5) \\ (4,5) \end{array} \begin{array}{c} \begin{array}{ccccc} 1 & 2 & 3 & 4 & 5 \end{array} \\ \begin{bmatrix} \mathbf{1} & \mathbf{-1} & 0 & 0 & 0 \\ 0 & \mathbf{1} & \mathbf{-1} & 0 & 0 \\ \mathbf{1} & 0 & 0 & \mathbf{-1} & 0 \\ 0 & \mathbf{1} & 0 & \mathbf{-1} & 0 \\ 0 & 0 & \mathbf{1} & 0 & \mathbf{-1} \\ 0 & 0 & 0 & \mathbf{1} & \mathbf{-1} \end{bmatrix} \end{array}$$

$$W(:,1) = \frac{X(:,1)+X(:,2)+X(:,3)}{3} = \frac{1}{3}\begin{pmatrix} 0 \\ 1 \\ 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}$$

$$W(:,2) = \frac{X(:,3)+X(:,4)}{2} = \frac{1}{2}\begin{pmatrix} 0 \\ -1 \\ -1 \\ -1 \\ 1 \\ 1 \end{pmatrix}$$

$$\begin{aligned}
&\textstyle\sum_{t=1}^{2} \|X(:,C_t) - W(:,t)\boldsymbol{e}^T\|_F^2 \\
&= \|X(:,C_1) - W(:,1)\boldsymbol{e}^T\|_F^2 + \|X(:,C_2) - W(:,2)\boldsymbol{e}^T\|_F^2 \\
&= \textstyle\sum_{e \notin \text{cut}} (1-0)^2 + (-1-0)^2 \\
&\quad + \textstyle\sum_{e \in \text{cut}} \frac{|C_1|-1}{|C_1|^2} + (1 - \frac{1}{|C_1|})^2 + \frac{|C_2|-1}{|C_2|^2} + (1 - \frac{1}{|C_2|})^2 \\
&= 2|E(C_1,C_1)| + 2|E(C_2,C_2)| + (2 - \frac{1}{|C_1|} - \frac{1}{|C_2|})|E(C_1,C_2)| \\
&= 2|E| - |V| \frac{|E(C_1,C_2)|}{|C_1|.|C_2|}.
\end{aligned}$$

$$\boxed{\min \textstyle\sum_{t=1}^{2} \|X(:,C_t) - W(:,t)\boldsymbol{e}^T\|_F^2 \iff \max \frac{|E(C_1,C_2)|}{|C_1|.|C_2|}.}$$

Plan

Approximation de rang 1 en norme de Frobenius

Soit $X \in \mathbb{R}^{m \times n}$,
$$\min_{w \in \mathbb{R}^m, h \in \mathbb{R}^n} \|X - wh^T\|_F^2.$$

- Résolvable en temps polynomial avec la SVD, alors que non-convexe !
  La solution optimale est $w = \alpha u_1$ et $h = \beta v_1$ with $\alpha\beta = \sigma_1$,
  où (σ_1, u_1, v_1) est le premier triplet singulier de X.

Brefs rappels

Etant donnée une matrice X de rang r, la SVD est : $X = U\Sigma V^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T$.

- $\|X\|_F^2 = \sigma_1^2 + ...\sigma_r^2$.
- Meilleure approx. de rang k : $X_k = \arg\min_{\tilde{X}} \|X - \tilde{X}\|_F^2$ s.t. $\operatorname{rank}(\tilde{X}) \leq k$,

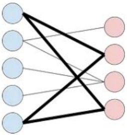
  solution : $X_k = \sum_{i=1}^{k} \sigma_i u_i v_i^T$ et $\|X - X_k\|_F^2 = \sigma_{k+1}^2 + ... + \sigma_r^2$.

Approximation de rang 1 avec $w \geq 0$, $h \geq 0$

R1-MF. Soit $X \in \mathbb{R}^{m \times n}$,

$$\min_{w \in \mathbb{R}^m_+, h \in \mathbb{R}^n_+} \|X - wh^T\|_F^2.$$

- NP-hard [Gillis, Glineur, 2008].
  Réduction du problème de biclique (la réduction uilise des entrées < 0).



$$M = \begin{pmatrix} 0 & \mathbf{1} & 0 & \mathbf{1} \\ 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & \mathbf{1} & 0 & \mathbf{1} \end{pmatrix} \qquad X = \begin{pmatrix} -d & 1 & -d & 1 \\ 1 & -d & 1 & -d \\ -d & -d & 1 & -d \\ -d & -d & 1 & -d \\ -d & 1 & -d & 1 \end{pmatrix}.$$

R1-NMF (cas particulier de R1-MF). Soit $X \in \mathbb{R}_+^{m \times n}$,

$$\min_{w \in \mathbb{R}^m_+, h \in \mathbb{R}^n_+} \|X - wh^T\|_F^2.$$

- Résolvable en temps polynomial avec la SVD : $w = \alpha |u_1|$ et $h = \beta |v_1|$ avec $\alpha\beta = \sigma_1$, où (σ_1, u_1, v_1) est le premier triplet singulier de X.
  En effet, $(X_{ij} - \sigma_1 u_{1i} v_{1j})^2 \geq (X_{ij} - \sigma_1 |u_{1i}||v_{1j}|)^2$ avec $X \geq 0$.

Complexité de la NMF en utilisant la norme de Frobenius

NMF. Etant donnés $X \in \mathbb{R}_+^{m \times n}$, $r \leq \min\{m, n\}$,

$$\min_{W \geq 0, H \geq 0} \|X - WH\|_F^2.$$

Exact NMF. Etant donnés $X \in \mathbb{R}_+^{m \times n}$, $r \leq \min\{m, n\}$,

$$\text{trouver } W \in \mathbb{R}_+^{m \times r},\ H \in \mathbb{R}_+^{r \times n} \text{ telles que } \quad X = WH.$$

NMF permet de résoudre Exact NMF, donc **NMF est aussi dur qu'Exact NMF**.

- Que peut-on dire sur la complexité de Exact NMF?

  - Vérifier que $\text{rank}_+(X) = r$ peut être effectué en $(mn)^{O(r^2)}$ [Moitra, 2013]
    MAIS, pas d'algorithme pratique jusqu'à présent (même pour un petit r)

- Et pour le problème NMF ?

  - NMF est NP-hard quand r n'est pas fixé (fait partie de l'énoncé)
  - Et quand r est fixé?

Problème de partitionnement : ReLU-NMD de rang 1

R1-ReLu-NMD.
Pour $X \in \mathbb{R}^{m \times n}$, on souhaite trouver $w^* \in \mathbb{R}^m$ et $h^* \in \mathbb{R}^n$ tels que

$$(w^*, h^*) = \arg \min_{w \in \mathbb{R}^m, h \in \mathbb{R}^n} \|X - \max(0, wh^T)\|_F^2.$$

$$\max(0, wh^\top) = \max\left(0, \begin{pmatrix} + \\ - \\ 0 \end{pmatrix} \begin{pmatrix} + \\ - \\ 0 \end{pmatrix}^\top\right) = \left(\begin{array}{c|c|c} + & 0 & 0 \\ \hline 0 & + & 0 \\ \hline 0 & 0 & 0 \end{array}\right).$$

Exemple : Pour $w = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$ et $h = \begin{pmatrix} 1 \\ -1 \end{pmatrix}$, on a :

$$\max(0, wh^\top) = \max\left(0, \begin{pmatrix} 1 \\ -1 \end{pmatrix} \begin{pmatrix} 1 \\ -1 \end{pmatrix}^T\right) = \max\left(0, \begin{pmatrix} 1 & -1 \\ -1 & 1 \end{pmatrix}\right) = \left(\begin{array}{c|c} 1 & 0 \\ \hline 0 & 1 \end{array}\right).$$

Complexité de ReLU-NMD de rang 1

On a vu dans les slides précédents, **R1-MF** est un problème NP-hard:

R1-MF. Given $R \in \mathbb{R}^{s \times t}$,
$$\min_{u \in \mathbb{R}^s_+, h \in \mathbb{R}^t_+} \|R - uv^T\|_F^2.$$

Que peut-on dire de la complexité de ce problème?

R1-ReLu-NMD. For $X \in \mathbb{R}^{m \times n}$, find $w^* \in \mathbb{R}^m$ and $h^* \in \mathbb{R}^n$ such that
$$(w^*, h^*) = \arg \min_{w \in \mathbb{R}^m, h \in \mathbb{R}^n} \|X - \max(0, wh^T)\|_F^2.$$

De toute instance R1-MF, on crée une instance de R1ReLuNMD avec $d > \|R\|_F$:
$$X = \left(\begin{array}{ccc|c} & & & -d \\ & R & & \vdots \\ & & & -d \\ \hline -d & \dots & -d & d \end{array} \right),$$

Idée de la preuve: à cause de l'entrée $= d$ en bas à droite $\Rightarrow w_m \neq 0 \Rightarrow$
$w_m = -1 \ \Rightarrow h_n < 0 \ \Rightarrow w_{1:m-1} \geq 0$ and $h_{1:n-1} \geq 0 \Rightarrow$ on résout R1-MF.
Problème ouvert : et dans le cas $X \geq 0$?

Component-Wise Squared Factorization

Component-Wise Squared Factorization (CSF):
Etant donnés une matrice $X \in \mathbb{R}_+^{m \times n}$ et un rang $r \in \mathbb{N}$,
on souhaite déterminer $W \in \mathbb{R}^{m \times r}$ et $H \in \mathbb{R}^{r \times n}$ résolvant le problème :

$$\min_{W,H} \|X - (WH)^{.2}\|_F^2.$$

Ce problème est NP-difficile.

PARTITION est un problème NP-hard :

INSTANCE: Un ensemble S d'entiers positifs s_i.

QUESTION: trouver une partition $S = S_1 \cup S_2$, $S_1 \cap S_2 = \emptyset$ telle que $\sum_{i \in S_1} s_i = \sum_{i \in S_2} s_i$.

Example: $S = \{3, 1, 2\}$.

De toute instance PARTITION, on crée une instance de CSF avec

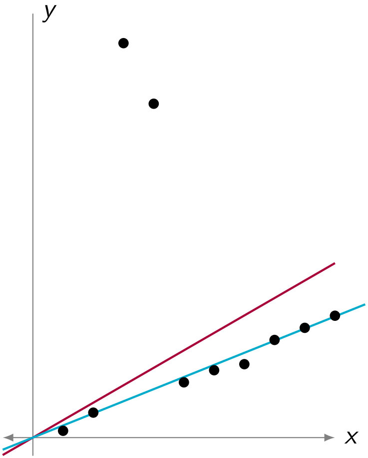
$$X = \begin{pmatrix} 1 & 0 & 0 & 9 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 4 \\ 1 & 1 & 1 & 0 \end{pmatrix}.$$

La dernière colonne contient les nombres s_i^2.

Plan

La norme ℓ_1 est plus robuste aux outliers

Etant donnés n points (x_i, y_i), on souhaite déterminer un modèle du type $y = \alpha x$.



Pour ℓ_2-norm, $\alpha^* = \arg\min_{\alpha} \ (y_1 - \alpha x_1)^2 + ... + (y_n - \alpha x_n)^2 \ = \arg\min_{\alpha} \|y - \alpha x\|_2$.

Pour ℓ_1-norm, $\alpha^* = \arg\min_{\alpha} |y_1 - \alpha x_1| + ... + |y_n - \alpha x_n| \ = \arg\min_{\alpha} \|y - \alpha x\|_1$.

Le modèle NMF utilisant la norme ℓ_1

Etant donnée une matrice $X \in \mathbb{R}^{m \times n}$,

$$\min_{W \geq 0, H \geq 0} \|X - WH\|_1 = \sum_{i=1}^{m} \sum_{j=1}^{n} |X - WH|_{ij}.$$

Exemples avec $r = 1$:

X	ℓ_2-NMF solution	ℓ_1-NMF solution
$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 \end{pmatrix}$	$\begin{pmatrix} 1.03 & 0.92 & 0.92 & 0.92 & 0.44 \\ 1.15 & 1.02 & 1.02 & 1.02 & 0.50 \\ 1.03 & 0.92 & 0.92 & 0.92 & 0.44 \\ 0.40 & 0.36 & 0.36 & 0.36 & 0.17 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$
$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}$	$\begin{pmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 4 \end{pmatrix}$	$\begin{pmatrix} 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix}$

- Ce modèle est plus robuste aux outliers
- Qu'en est-il de la complexité ?

Complexité du modèle ℓ_1-MF dans le cas de rang 1

[Gillis, Vavasis, 2015] ont prouvé que le problème est difficile dans le cas où X est uniquement composée de ± 1 :

Given $X \in \{\pm 1\}^{m \times n}$, $$\min_{w \in \mathbb{R}^m, h \in \mathbb{R}^n} \|X - wh^T\|_1.$$

La preuve peut être structurée en trois grandes étapes :

Step 1.
$$\min_{w \in \mathbb{R}^m, h \in \mathbb{R}^n} \|X - wh^T\|_1 = \min_{w \in \{\pm 1\}^m, h \in \{\pm 1\}^n} \|X - wh^T\|_1.$$

Step 2.
$$\min_{w \in \{\pm 1\}^m, h \in \{\pm 1\}^n} \|X - wh^T\|_1 \text{ est équivalent à } \max_{w \in \{\pm 1\}^m, h \in \{\pm 1\}^n} w^T X h.$$

Step 3.
$$\text{résoudre } \max_{w \in \{\pm 1\}^m, h \in \{\pm 1\}^n} w^T X h \quad \text{est NP-hard.}$$

Problèmes ouverts : et dans le cas $X \geq 0$? et dans le cas $X \in \{0, 1\}^{m \times n}$?