Cours 13 : Méthodes de gradient

Christophe Gonzales

LIP6 – Université Paris 6, France

Méthodes primales / duales

- $f : \mathbb{R}^n \mapsto \mathbb{R}$
- $\min f(x)$
  $s.c.\ \ g_i(x) \leq 0,\ i = \{1, \ldots, m\}, \quad x \in \mathbb{R}^n$

Méthodes primales / duales

- *méthodes primales :*
  - génération d'une séquence de solutions, i.e., de points satisfaisant les contraintes
  - séquence $\Longrightarrow$ fait décroître f $\Longrightarrow$ propriété anytime
  - méthodes difficiles à implanter et à calibrer

Méthodes primales / duales

- $f : \mathbb{R}^n \mapsto \mathbb{R}$
- $\min f(x)$
  $s.c.\ \ g_i(x) \leq 0,\ i = \{1, \dots, m\}, \quad x \in \mathbb{R}^n$

Méthodes primales / duales

- *méthodes primales :*
  - génération d'une séquence de solutions, i.e., de points satisfaisant les contraintes
  - séquence $\Longrightarrow$ fait décroître f $\Longrightarrow$ propriété anytime
  - méthodes difficiles à implanter et à calibrer
- *méthodes duales :*
  - résolution d'une séquence de problèmes d'optimisation sans contraintes $\Longrightarrow$ utilisation du lagrangien
  - méthodes plus robustes que les primales et convergence plus facile à assurer
  - pas de propriété anytime

⚠ si $\min_{x \in \mathbb{R}^n} f(x)$ obtenu pour $\widehat{x}$ tel que $g_i(\widehat{x}) < 0 \ \forall i$
alors $\widehat{x}$ optimum du problème sans contraintes

Méthode du gradient projeté

si $\min_{x \in \mathbb{R}^n} f(x)$ obtenu pour $\widehat{x}$ tel que $g_i(\widehat{x}) < 0 \; \forall i$
alors $\widehat{x}$ optimum du problème sans contraintes

Sinon le min de $f(x)$ sous contraintes obtenu pour $\widehat{x}$ sur la frontière des réalisables

Méthode du gradient projeté

si $\min_{x \in \mathbb{R}^n} f(x)$ obtenu pour $\widehat{x}$ tel que $g_i(\widehat{x}) < 0 \ \forall i$
alors $\widehat{x}$ optimum du problème sans contraintes

Sinon le min de $f(x)$ sous contraintes obtenu pour $\widehat{x}$ sur la frontière des réalisables

Méthode du gradient projeté = méthode primale

Méthode du gradient projeté

Principe : adapter les méthodes de gradient des problèmes d'optimisation sans contrainte

Méthode du gradient projeté

⚠ si $\min_{x \in \mathbb{R}^n} f(x)$ obtenu pour $\widehat{x}$ tel que $g_i(\widehat{x}) < 0 \; \forall i$
alors $\widehat{x}$ optimum du problème sans contraintes

Sinon le min de $f(x)$ sous contraintes obtenu pour $\widehat{x}$ sur la frontière des réalisables

Méthode du gradient projeté = méthode primale

Méthode du gradient projeté

Principe : adapter les méthodes de gradient des problèmes d'optimisation sans contrainte

$\Longrightarrow$ projeter les déplacements sur la frontière de la région des réalisables

Méthode du gradient projeté

⚠ si $\min_{x \in \mathbb{R}^n} f(x)$ obtenu pour $\widehat{x}$ tel que $g_i(\widehat{x}) < 0 \ \forall i$
alors $\widehat{x}$ optimum du problème sans contraintes

Sinon le min de $f(x)$ sous contraintes obtenu pour $\widehat{x}$ sur la frontière des réalisables

Méthode du gradient projeté = méthode primale

Méthode du gradient projeté

Principe : adapter les méthodes de gradient des problèmes d'optimisation sans contrainte

$\Longrightarrow$ projeter les déplacements sur la frontière de la région des réalisables

$\Longrightarrow$ génération d'une séquence de points réalisables

Méthode du gradient projeté

⚠ si $\min_{x \in \mathbb{R}^n} f(x)$ obtenu pour $\widehat{x}$ tel que $g_i(\widehat{x}) < 0 \ \forall i$
alors $\widehat{x}$ optimum du problème sans contraintes

Sinon le min de $f(x)$ sous contraintes obtenu pour $\widehat{x}$ sur la frontière des réalisables

Méthode du gradient projeté = méthode primale

Méthode du gradient projeté

Principe : adapter les méthodes de gradient des problèmes d'optimisation sans contrainte

$\Longrightarrow$ projeter les déplacements sur la frontière de la région des réalisables

$\Longrightarrow$ génération d'une séquence de points réalisables

Exemple : méthode de Rosen (1960)

Méthode du gradient projeté de Rosen

Idée : projeter le gradient sur la frontière des réalisables
$\Longrightarrow$ chemin sur la frontière dans la direction de la plus grande pente

Méthode du gradient projeté de Rosen

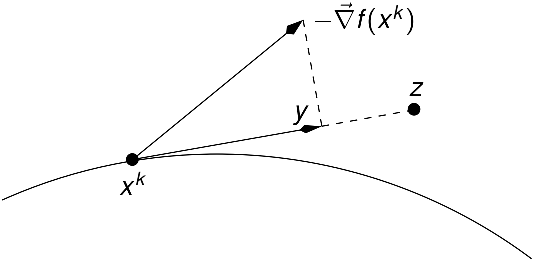
Idée : projeter le gradient sur la frontière des réalisables
$\Longrightarrow$ chemin sur la frontière dans la direction de la plus grande pente

$\Longrightarrow$ fonctionne essentiellement avec de contraintes linéaires

Méthode du gradient projeté de Rosen

Idée : projeter le gradient sur la frontière des réalisables
$\Longrightarrow$ chemin sur la frontière dans la direction de la plus grande pente

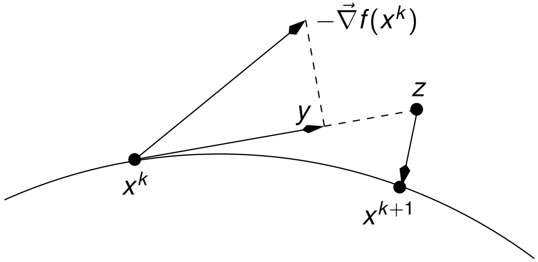
$\Longrightarrow$ fonctionne essentiellement avec de contraintes linéaires :



Méthode du gradient projeté de Rosen

Idée : projeter le gradient sur la frontière des réalisables
$\Longrightarrow$ chemin sur la frontière dans la direction de la plus grande pente

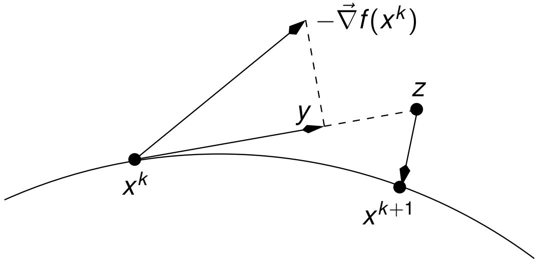
$\Longrightarrow$ fonctionne essentiellement avec de contraintes linéaires :



Méthode du gradient projeté de Rosen

Idée : projeter le gradient sur la frontière des réalisables
$\Longrightarrow$ chemin sur la frontière dans la direction de la plus grande pente

$\Longrightarrow$ fonctionne essentiellement avec de contraintes linéaires :



contraintes linéaires $\Longrightarrow$ le déplacement reste sur l'ensemble des réalisables

Problème non linéaire avec contraintes linéaires

$$(P) \qquad \begin{array}{l} \min f(x) \\ s.c. \begin{cases} a_i^T x \leq b_i, & i \in I_1 \\ a_i^T x = b_i, & i \in I_2 \end{cases} \end{array}$$

Remarques :

- pas de contraintes de signes sur les $x_i \Longrightarrow$ incluses dans I_1
- les contraintes de I_2 peuvent être transformées en I_1 (dédoublement)

$$(P) \qquad \begin{array}{l} \min f(x) \\ s.c. \begin{cases} a_i^T x \leq b_i, & i \in I_1 \\ a_i^T x = b_i, & i \in I_2 \end{cases} \end{array}$$

Remarques :

- pas de contraintes de signes sur les $x_i \Longrightarrow$ incluses dans I_1
- les contraintes de I_2 peuvent être transformées en I_1 (dédoublement)

Prochains slides : résolution de (P) par la méthode de Rosen

Résolution de (P) par la méthode de Rosen (1/9)

$$(P) \qquad \begin{array}{l} \min f(x) \\ s.c. \begin{cases} a_i^T x \leq b_i, & i \in I_1 \\ a_i^T x = b_i, & i \in I_2 \end{cases} \end{array}$$

- On suppose qu'on connaît une solution x

  $\Longrightarrow$ au pire, utiliser un simplexe pour trouver un x

Résolution de (P) par la méthode de Rosen (1/9)

$$(P) \qquad \begin{array}{l} \min f(x) \\ s.c. \begin{cases} a_i^T x \leq b_i, & i \in I_1 \\ a_i^T x = b_i, & i \in I_2 \end{cases} \end{array}$$

- On suppose qu'on connaît une solution x
  $\Longrightarrow$ au pire, utiliser un simplexe pour trouver un x
- $I^0 = \{i \in I_1 : a_i^T x = b_i\} \cup I_2$
  $=$ ensemble des contraintes serrées

Résolution de (P) par la méthode de Rosen (1/9)

$$(P) \qquad \begin{array}{l} \min f(x) \\ s.c. \begin{cases} a_i^T x \leq b_i, & i \in I_1 \\ a_i^T x = b_i, & i \in I_2 \end{cases} \end{array}$$

- On suppose qu'on connaît une solution x
  $\Longrightarrow$ au pire, utiliser un simplexe pour trouver un x
- $I^0 = \{i \in I_1 : a_i^T x = b_i\} \cup I_2$
  $=$ ensemble des contraintes serrées
- $A^0 =$ matrice composée des lignes a_i^T, $i \in I^0$

Résolution de (P) par la méthode de Rosen (1/9)

$$(P) \qquad \begin{array}{l} \min f(x) \\ s.c. \begin{cases} a_i^T x \leq b_i, & i \in I_1 \\ a_i^T x = b_i, & i \in I_2 \end{cases} \end{array}$$

- On suppose qu'on connaît une solution x
  $\Longrightarrow$ au pire, utiliser un simplexe pour trouver un x
- $I^0 = \{i \in I_1 : a_i^T x = b_i\} \cup I_2$
  $=$ ensemble des contraintes serrées
- $A^0 =$ matrice composée des lignes a_i^T, $i \in I^0$
- Rosen $\Longrightarrow$ déplacements réalisables
  $\Longrightarrow$ si d est une direction alors $a_i^T d = 0 \ \forall i \in I^0$
  $\Longrightarrow A^0 d = 0$

$$\text{Déplacement } d \Longrightarrow \text{faire décroître } f$$
$$\Longrightarrow \text{minimiser } \vec{\nabla} f(x)^T d$$

Résolution de (P) par la méthode de Rosen (2/9)

Déplacement $d \Longrightarrow$ faire décroître f

$\Longrightarrow$ minimiser $\vec{\nabla} f(x)^T d$

Principe général de la méthode

1. partir d'un point x réalisable
2. calculer une direction d solution de :

$$\min \vec{\nabla} f(x)^T d$$

$$s.c. \begin{cases} A^0 d = 0 \\ \|d\| = 1 \end{cases}$$

3. se déplacer selon d pour obtenir un nouvel x
4. si l'on n'est pas sur l'optimum, revenir en 1

Déplacement $d \Longrightarrow$ faire décroître f

$\Longrightarrow$ minimiser $\vec{\nabla} f(x)^T d$

Principe général de la méthode

1. partir d'un point x réalisable
2. calculer une direction d solution de :

$$\min \vec{\nabla} f(x)^T d$$
$$s.c. \begin{cases} A^0 d = 0 \\ \|d\| = 1 \end{cases}$$

3. se déplacer selon d pour obtenir un nouvel x
4. si l'on n'est pas sur l'optimum, revenir en 1

Problème : peut-on résoudre ce problème rapidement ?

Théorème

- A^0 matrice $q \times n$, $q \leq n$, de rang q

  $\Longrightarrow$ les a_i^T sont linéairement indépendants

  (pas de dégénérescence)

- la solution optimale du problème :

$$\min \vec{\nabla} f(x)^T d$$

$$s.c. \begin{cases} A^0 d = 0 \\ \|d\| = 1 \end{cases}$$

est $d = \overline{y}/\|\overline{y}\|$, où $\overline{y}$ = projection de $-\vec{\nabla} f(x)$ sur $\{y : A^0 y = 0\}$

Théorème

- A^0 matrice $q \times n$, $q \leq n$, de rang q
  $\Longrightarrow$ les a_i^T sont linéairement indépendants
  (pas de dégénérescence)
- la solution optimale du problème :

$$\min \vec{\nabla} f(x)^T d$$

$$s.c. \begin{cases} A^0 d = 0 \\ \|d\| = 1 \end{cases}$$

est $d = \overline{y}/\|\overline{y}\|$, où $\overline{y}$ = projection de $-\vec{\nabla} f(x)$ sur $\{y : A^0 y = 0\}$

- $\overline{y} = -P^0 \vec{\nabla} f(x) = -\left(Id - A^{0T} \left[A^0 A^{0T}\right]^{-1} A^0 \right) \vec{\nabla} f(x)$

Théorème

- A^0 matrice $q \times n$, $q \leq n$, de rang q
  $\Longrightarrow$ les a_i^T sont linéairement indépendants
  (pas de dégénérescence)
- la solution optimale du problème :

$$\min \vec{\nabla} f(x)^T d$$
$$s.c. \begin{cases} A^0 d = 0 \\ \|d\| = 1 \end{cases}$$

est $d = \overline{y}/\|\overline{y}\|$, où $\overline{y}$ = projection de $-\vec{\nabla} f(x)$ sur $\{y : A^0 y = 0\}$

- $\overline{y} = -P^0 \vec{\nabla} f(x) = -\left(Id - A^{0T}\left[A^0 A^{0T}\right]^{-1} A^0\right) \vec{\nabla} f(x)$

  $P^0 =$ matrice de projection sur $\{y : A^0 y = 0\}$

  $Id =$ matrice identité

Résolution de (P) par la méthode de Rosen (4/9)

Démonstration :

soit $S^0 = \{y : A^0 y = 0\}$

soit $S^{0\perp}$ = sous-espace orthogonal à S^0 dans $\mathbb{R}^n$

Résolution de (P) par la méthode de Rosen (4/9)

Démonstration :

soit $S^0 = \{y : A^0 y = 0\}$

soit $S^{0\perp}$ = sous-espace orthogonal à S^0 dans $\mathbb{R}^n$

$\forall x \in \mathbb{R}^n,\ \exists \overline{y} \in S^0$ et $\exists \overline{z} \in S^{0\perp}$ tels que $x = \overline{y} + \overline{z}$

Démonstration :

soit $S^0 = \{y : A^0 y = 0\}$

soit $S^{0\perp}$ = sous-espace orthogonal à S^0 dans $\mathbb{R}^n$

$\forall x \in \mathbb{R}^n, \ \exists \overline{y} \in S^0$ et $\exists \overline{z} \in S^{0\perp}$ tels que $x = \overline{y} + \overline{z}$

$\Longrightarrow$ en particulier $-\vec{\nabla} f(x) = \overline{y} + \overline{z}$

$\Longrightarrow -\vec{\nabla} f(x)^T d = \overline{y}^T d + \overline{z}^T d$

Démonstration :

soit $S^0 = \{y : A^0 y = 0\}$

soit $S^{0\perp}$ = sous-espace orthogonal à S^0 dans $\mathbb{R}^n$

$\forall x \in \mathbb{R}^n,\ \exists \overline{y} \in S^0$ et $\exists \overline{z} \in S^{0\perp}$ tels que $x = \overline{y} + \overline{z}$

$\Longrightarrow$ en particulier $-\vec{\nabla} f(x) = \overline{y} + \overline{z}$

$\Longrightarrow -\vec{\nabla} f(x)^T d = \overline{y}^T d + \overline{z}^T d$

Or $d \in S^0 \Longrightarrow \overline{z}^T d = 0 \Longrightarrow -\vec{\nabla} f(x)^T d = \overline{y}^T d$

$\Longrightarrow \vec{\nabla} f(x)^T d$ minimal quand $\overline{y}^T d$ maximal

Démonstration :

soit $S^0 = \{y : A^0 y = 0\}$

soit $S^{0\perp}$ = sous-espace orthogonal à S^0 dans $\mathbb{R}^n$

$\forall x \in \mathbb{R}^n,\ \exists \overline{y} \in S^0$ et $\exists \overline{z} \in S^{0\perp}$ tels que $x = \overline{y} + \overline{z}$

$\Longrightarrow$ en particulier $-\vec{\nabla} f(x) = \overline{y} + \overline{z}$

$\Longrightarrow -\vec{\nabla} f(x)^T d = \overline{y}^T d + \overline{z}^T d$

Or $d \in S^0 \Longrightarrow \overline{z}^T d = 0 \Longrightarrow -\vec{\nabla} f(x)^T d = \overline{y}^T d$

$\Longrightarrow \vec{\nabla} f(x)^T d$ minimal quand $\overline{y}^T d$ maximal

$d \in S^0$, $\|d\| = 1$ et $\overline{y} \in S^0 \Longrightarrow \overline{y}^T d$ maximal pour $d = \overline{y}/\|\overline{y}\|$

$\Longrightarrow$ il reste maintenant à trouver l'expression de $\overline{y}$

caractérisation de $\overline{y}$:

remarque : $S^{0\perp}$ est le sous-espace de $\mathbb{R}^n$ généré par les colonnes de $A^{0\,T}$

$\Longrightarrow \forall \overline{z} \in S^{0\perp}, \exists u$ tel que $\overline{z} = A^{0\,T} u$

caractérisation de $\overline{y}$:

remarque : $S^{0\perp}$ est le sous-espace de $\mathbb{R}^n$ généré par les colonnes de A^{0T}

$\Longrightarrow \forall \overline{z} \in S^{0\perp}, \exists u$ tel que $\overline{z} = A^{0T} u$

$\Longrightarrow \exists u$ tel que $-\vec{\nabla} f(x) = \overline{y} + A^{0T} u$

caractérisation de $\overline{y}$:

remarque : $S^{0\perp}$ est le sous-espace de $\mathbb{R}^n$ généré par les colonnes de $A^{0\,T}$

$\Longrightarrow \forall \overline{z} \in S^{0\perp}, \exists u$ tel que $\overline{z} = A^{0\,T} u$

$\Longrightarrow \exists u$ tel que $-\vec{\nabla} f(x) = \overline{y} + A^{0\,T} u$

$\overline{y} \in S^0 \Longrightarrow A^0 \overline{y} = 0$

caractérisation de $\overline{y}$:

remarque : $S^{0\perp}$ est le sous-espace de $\mathbb{R}^n$ généré par les colonnes de $A^{0\,T}$

$\Longrightarrow \forall \overline{z} \in S^{0\perp}, \exists u$ tel que $\overline{z} = A^{0\,T} u$

$\Longrightarrow \exists u$ tel que $-\vec{\nabla} f(x) = \overline{y} + A^{0\,T} u$

$\overline{y} \in S^0 \Longrightarrow A^0 \overline{y} = 0 \Longrightarrow A^0 \overline{y} = -A^0 \vec{\nabla} f(x) - A^0 A^{0\,T} u = 0$

caractérisation de $\overline{y}$:

remarque : $S^{0\perp}$ est le sous-espace de $\mathbb{R}^n$ généré par les colonnes de A^{0T}

$\Longrightarrow \forall \overline{z} \in S^{0\perp}, \exists u$ tel que $\overline{z} = A^{0T}u$

$\Longrightarrow \exists u$ tel que $-\vec{\nabla} f(x) = \overline{y} + A^{0T}u$

$\overline{y} \in S^0 \Longrightarrow A^0\overline{y} = 0 \Longrightarrow A^0\overline{y} = -A^0\vec{\nabla} f(x) - A^0A^{0T}u = 0$

Or A^0 de rang $q \Longrightarrow A^0A^{0T}$ inversible

$\Longrightarrow u = -\left[A^0A^{0T}\right]^{-1} A^0\vec{\nabla} f(x)$

caractérisation de $\overline{y}$:

remarque : $S^{0\perp}$ est le sous-espace de $\mathbb{R}^n$ généré par les colonnes de A^{0T}

$\Longrightarrow \forall \overline{z} \in S^{0\perp}, \exists u$ tel que $\overline{z} = A^{0T} u$

$\Longrightarrow \exists u$ tel que $-\vec{\nabla} f(x) = \overline{y} + A^{0T} u$

$\overline{y} \in S^0 \Longrightarrow A^0 \overline{y} = 0 \Longrightarrow A^0 \overline{y} = -A^0 \vec{\nabla} f(x) - A^0 A^{0T} u = 0$

Or A^0 de rang $q \Longrightarrow A^0 A^{0T}$ inversible

$\Longrightarrow u = -\left[A^0 A^{0T}\right]^{-1} A^0 \vec{\nabla} f(x)$

$$\overline{y} = -\vec{\nabla} f(x) - A^{0T} u = -\vec{\nabla} f(x) + A^{0T} \left[A^0 A^{0T}\right]^{-1} A^0 \vec{\nabla} f(x)$$

$$= -\left(Id - A^{0T} \left[A^0 A^{0T}\right]^{-1} A^0\right) \vec{\nabla} f(x)$$

CQFD

Rappel du théorème

La solution optimale du problème :

$$\begin{aligned} &\min \vec{\nabla} f(x)^T d \\ &s.c. \quad A^0 d = 0, \quad \|d\| = 1 \end{aligned}$$

est $d = \overline{y}/\|\overline{y}\|$, où $\overline{y}$ = projection de $-\vec{\nabla} f(x)$ sur $\{y : A^0 y = 0\}$

et $\overline{y} = -\left(Id - A^{0T}\left[A^0 A^{0T}\right]^{-1} A^0\right)\vec{\nabla} f(x)$

si $\overline{y} \neq 0$ alors $\overline{y}$ est une direction de descente de gradient :

$\vec{\nabla} f(x)^T \overline{y} = -(\overline{y}^T + \overline{z}^T)\overline{y} = -\overline{y}^T \overline{y} < 0$

Résolution de (P) par la méthode de Rosen (7/9)

Rappel de la méthode de Rosen

1. partir d'un point x réalisable
2. calculer la direction $d = \overline{y}/\|\overline{y}\|$
3. se déplacer selon d pour obtenir un nouvel x réalisable
4. si l'on n'est pas sur l'optimum, revenir en 1

Rappel de la méthode de Rosen

1. partir d'un point x réalisable
2. calculer la direction $d = \overline{y}/\|\overline{y}\|$
3. se déplacer selon d pour obtenir un nouvel x réalisable
4. si l'on n'est pas sur l'optimum, revenir en 1

on peut calculer la longueur de déplacement max selon d :

$$\alpha_{\max} = \max\{\alpha \geq 0 : x + \alpha\overline{y} \in X\}$$

$$\text{où } X = \{x \in \mathbb{R}^n : a_i x \leq b_i \ \forall i \in I_1 \text{ et } a_i x = b_i \ \forall i \in I_2\}$$

Rappel de la méthode de Rosen

1. partir d'un point x réalisable
2. calculer la direction $d = \overline{y}/\|\overline{y}\|$
3. se déplacer selon d pour obtenir un nouvel x réalisable
4. si l'on n'est pas sur l'optimum, revenir en 1

on peut calculer la longueur de déplacement max selon d :

$$\alpha_{\max} = \max\{\alpha \geq 0 : x + \alpha\overline{y} \in X\}$$

$$\text{où } X = \{x \in \mathbb{R}^n : a_i x \leq b_i \ \forall i \in I_1 \text{ et } a_i x = b_i \ \forall i \in I_2\}$$

prochain point réalisable : $x' = x + \overline{\alpha}\overline{y}$ avec $\overline{\alpha} = \min\limits_{\alpha \in [0, \alpha_{\max}]} f(x + \alpha\overline{y})$

Rappel de la méthode de Rosen

1. partir d'un point x réalisable
2. calculer la direction $d = \overline{y}/\|\overline{y}\|$
3. se déplacer selon d pour obtenir un nouvel x réalisable
4. si l'on n'est pas sur l'optimum, revenir en 1

on peut calculer la longueur de déplacement max selon d :

$$\alpha_{\text{max}} = \max\{\alpha \geq 0 : x + \alpha\overline{y} \in X\}$$

$$\text{où } X = \{x \in \mathbb{R}^n : a_i x \leq b_i \ \forall i \in I_1 \text{ et } a_i x = b_i \ \forall i \in I_2\}$$

prochain point réalisable : $x' = x + \overline{\alpha}\overline{y}$ avec $\overline{\alpha} = \min_{\alpha\in[0,\alpha_{\text{max}}]} f(x + \alpha\overline{y})$

en x', recalculer la nouvelle matrice de projection, etc

Problème : et si $\overline{y} = 0$?

Problème : et si $\overline{y} = 0$?

$$\overline{y} = -\vec{\nabla} f(x) - {A^0}^T u$$

$$\Longrightarrow \begin{cases} \vec{\nabla} f(x) + {A^0}^T u = 0 \\ u = -\left[A^0 {A^0}^T\right]^{-1} A^0 \vec{\nabla} f(x) \end{cases}$$

Si $u \geq 0$, ça vous rappelle quelque chose ?

Problème : et si $\overline{y} = 0$?

$$\overline{y} = -\vec{\nabla} f(x) - A^{0T} u$$

$$\Longrightarrow \begin{cases} \vec{\nabla} f(x) + A^{0T} u = 0 \\ u = -\left[A^0 A^{0T}\right]^{-1} A^0 \vec{\nabla} f(x) \end{cases}$$

Si $u \geq 0$, ça vous rappelle quelque chose ?

damned, mais ce sont les conditions de Kuhn et Tucker ! ! ! ! !

Rappel : conditions de Kuhn et Tucker

condition nécessaire pour que $\widehat{x}$ = minimum local de f :

$\exists \lambda_i \geq 0$, $i = 1, \dots, m$, tels que :

1. $\vec{\nabla} f(\widehat{x}) + \sum_{i=1}^{m} \lambda_i \vec{\nabla} g_i(\widehat{x}) = 0$
2. $\lambda_i g_i(\widehat{x}) = 0, \quad \forall i \in \{1, \dots, m\}$

Donc si $\overline{y} = 0$ et $u \geq 0$:

Kuhn et Tucker $\Longrightarrow$ x est un optimum local

Donc si $\overline{y} = 0$ et $u \geq 0$:

Kuhn et Tucker $\Longrightarrow$ x est un optimum local

si $\overline{y} = 0$ et u a des composants strictement négatifs :

$\Longrightarrow$ x n'est pas optimal (Kuhn et Tucker = cond nécessaire)

$\Longrightarrow$ trouver une nouvelle direction d

Donc si $\overline{y} = 0$ et $u \geq 0$:

Kuhn et Tucker $\Longrightarrow$ x est un optimum local

si $\overline{y} = 0$ et u a des composants strictement négatifs :

$\Longrightarrow$ x n'est pas optimal (Kuhn et Tucker = cond nécessaire)

$\Longrightarrow$ trouver une nouvelle direction d

supprimer de A^0 une ligne a_i^T pour laquelle $u_i < 0$

$\Longrightarrow$ nouvelle matrice de projection

$\Longrightarrow$ nouvelle direction $\overline{y'}$ telle que $\overline{y'} \neq 0$ ($\Longrightarrow$ descente de gradient)

Donc si $\overline{y} = 0$ et $u \geq 0$:

Kuhn et Tucker $\Longrightarrow$ x est un optimum local

si $\overline{y} = 0$ et u a des composants strictement négatifs :

$\Longrightarrow$ x n'est pas optimal (Kuhn et Tucker = cond nécessaire)

$\Longrightarrow$ trouver une nouvelle direction d

supprimer de A^0 une ligne a_i^T pour laquelle $u_i < 0$

$\Longrightarrow$ nouvelle matrice de projection

$\Longrightarrow$ nouvelle direction $\overline{y'}$ telle que $\overline{y'} \neq 0$ ($\Longrightarrow$ descente de gradient)

choisir par exemple la ligne a_i^T pour laquelle u_i min

$\Longrightarrow$ complète la méthode du gradient de Rosen

Méthode du gradient projeté de Rosen

Algorithme de Rosen

1. étape $k = 0$: on part du point x^0
2. étape k : on est sur le point x^k
   Trouver l'ensemble $I^0(x^k)$ des contraintes serrées
   Soit $L^0 = I^0(x^k)$
3. Soit A^0 la matrice des lignes a_i^T, $i \in L^0$
   calculer $P^0 = Id - A^{0T}\left[A^0 A^{0T}\right]^{-1} A^0$
   calculer $y^k = -P^0 \vec{\nabla} f(x^k)$
   si $y^k = 0$ alors aller en 5
4. $y^k \neq 0$: calculer $\alpha_{\max} = \max\{\alpha \geq 0 : x^k + \alpha y^k \in X\}$
   calculer x^{k+1} tel que $f(x^{k+1}) = \min_{\alpha \in [0, \alpha_{\max}]}\{f(x^k + \alpha y^k)\}$
   $k \leftarrow k + 1$; retourner en 2
5. calculer $u = -\left[A^0 A^{0T}\right]^{-1} A^0 \vec{\nabla} f(x)$
   si $u \geq 0$ alors x^k optimum (local)
   sinon soit u_i l'élément le plus négatif de u
   $L^0 \leftarrow L^0 - \{i\}$; retourner en 3