Ploblèmes relaxés du sac-à-dos stochastique

Stefanie Kosuch and Abdel Lisser

Université Paris XI - Sud
LRI - GraphComb

02.02.2009

Le déroulement de la présentation

Le déroulement de la présentation

Le déroulement de la présentation

1. Introduction
2. Le problème du sac-à-dos stochastique
   - Formulation mathématique
   - Applications
3. Méthode de résolution
   - Ploblème relaxé
   - Approximation par convolution
   - Intégration par parties

Le déroulement de la présentation

Le problème du sac-à-dos déterministe

Le problème du sac-à-dos déterministe

- $c > 0$: capacité du sac-à-dos

Le problème du sac-à-dos déterministe

- $c > 0$: capacité du sac-à-dos
- n objets

Le problème du sac-à-dos déterministe

- $c > 0$: capacité du sac-à-dos
- n objets
- $r_i > 0$: bénéfice par unité de poids de l'objet i

Le problème du sac-à-dos déterministe

- $c > 0$: capacité du sac-à-dos
- n objets
- $r_i > 0$: bénéfice par unité de poids de l'objet i
- w_i : poids de l'objet i

Le problème du sac-à-dos déterministe

- $c > 0$: capacité du sac-à-dos
- n objets
- $r_i > 0$: bénéfice par unité de poids de l'objet i
- w_i : poids de l'objet i

- $x \in \{0,1\}^n$: vecteur des variables de décision

Le problème du sac-à-dos stochastique avec poids aléatoires

- $c > 0$: capacité du sac-à-dos
- n objets
- $r_i > 0$: bénéfice par unité de poids de l'objet i
- χ_i : poids de l'objet i, distribué normalement et indépendent
- $x \in \{0, 1\}^n$: vecteur des variables de décision

Le problème du sac-à-dos stochastique avec poids aléatoires

- $c > 0$: capacité du sac-à-dos
- n objets
- $r_i > 0$: bénéfice par unité de poids de l'objet i
- χ_i : poids de l'objet i, distribué normalement et indépendent
- μ_i,σ_i : moyenne et écart-type de χ_i
- $x \in \{0, 1\}^n$: vecteur des variables de décision

Outline

Problème du sac-à-dos avec recours simple (SRKP)

$$\max_{x\in\{0,1\}^n} \mathbb{E}[\sum_{i=1}^{n} r_i\chi_i x_i] - d\cdot\mathbb{E}[[\sum_{i=1}^{n}\chi_i x_i - c]^+]$$

Problème du sac-à-dos avec recours simple (SRKP)

$$\max_{x\in\{0,1\}^n} \mathbb{E}[\sum_{i=1}^{n} r_i\chi_i x_i] - d \cdot \mathbb{E}[[\sum_{i=1}^{n} \chi_i x_i - c]^+]$$

- $[x]^+ := \max(0, x) = x \cdot \mathbb{1}_{\mathbb{R}^+}(x)$ $(x \in \mathbb{R})$

Problème du sac-à-dos avec recours simple (SRKP)

$$\max_{x \in \{0,1\}^n} \mathbb{E}[\sum_{i=1}^{n} r_i \chi_i x_i] - d \cdot \mathbb{E}[[\sum_{i=1}^{n} \chi_i x_i - c]^+]$$

- $[x]^+ := \max(0, x) = x \cdot \mathbb{1}_{\mathbb{R}^+}(x)$ $(x \in \mathbb{R})$
- $\mathbb{1}_{\mathbb{R}^+}$: fonction indicatrice de l'interval réel non-négatif

Problème du sac-à-dos avec recours simple (SRKP)

$$\max_{x \in \{0,1\}^n} \mathbb{E}[\sum_{i=1}^{n} r_i \chi_i x_i] - d \cdot \mathbb{E}[[\sum_{i=1}^{n} \chi_i x_i - c]^+]$$

- $[x]^+ := \max(0, x) = x \cdot \mathbb{1}_{\mathbb{R}^+}(x)$ $(x \in \mathbb{R})$
- $\mathbb{1}_{\mathbb{R}^+}$: fonction indicatrice de l'interval réel non-négatif
- $d > 0$

Problème du sac-à-dos avec recours simple (SRKP)

$$\max_{x\in\{0,1\}^n} \mathbb{E}[\sum_{i=1}^{n} r_i\chi_i x_i] - d\cdot\mathbb{E}[[\sum_{i=1}^{n}\chi_i x_i - c]^+]$$

- $[x]^+ := \max(0, x) = x\cdot\mathbb{1}_{\mathbb{R}^+}(x)$ $(x\in\mathbb{R})$
- $\mathbb{1}_{\mathbb{R}^+}$: fonction indicatrice de l'interval réel non-négatif
- $d > 0$

Problème du sac-à-dos avec recours simple (SRKP)

$$\max_{x \in \{0,1\}^n} \mathbb{E}[\sum_{i=1}^{n} r_i \chi_i x_i] - d \cdot \mathbb{E}[[\sum_{i=1}^{n} \chi_i x_i - c]^+]$$

- $[x]^+ := \max(0, x) = x \cdot \mathbb{1}_{\mathbb{R}^+}(x)$ $(x \in \mathbb{R})$
- $\mathbb{1}_{\mathbb{R}^+}$: fonction indicatrice de l'interval réel non-négatif
- $d > 0$: facteur de pénalité par unité de surcharge

Problème du sac-à-dos avec contrainte en espérance (ECKP)

$$\max_{x \in \{0,1\}^n} \quad \mathbb{E}[\sum_{i=1}^{n} r_i \chi_i x_i]$$

$$\text{s.c.} \quad \mathbb{E}[\mathbb{1}_{\mathbb{R}^+}(c - \sum_{i=1}^{n} \chi_i x_i)] \geq p$$

Problème du sac-à-dos avec contrainte en espérance (ECKP)

$$\max_{x \in \{0,1\}^n} \quad \mathbb{E}[\sum_{i=1}^{n} r_i \chi_i x_i]$$

$$\text{s.c.} \quad \mathbb{E}[\mathbb{1}_{\mathbb{R}^+}(c - \sum_{i=1}^{n} \chi_i x_i)] \geq p$$

- $\mathbb{1}_{\mathbb{R}^+}$: fonction indicatrice de l'interval réel non-négatif

Problème du sac-à-dos avec contrainte en espérance (ECKP)

$$\max_{x \in \{0,1\}^n} \quad \mathbb{E}[\sum_{i=1}^{n} r_i \chi_i x_i]$$

$$\text{s.c.} \quad \mathbb{E}[\mathbb{1}_{\mathbb{R}^+}(c - \sum_{i=1}^{n} \chi_i x_i)] \geq p$$

- $\mathbb{1}_{\mathbb{R}^+}$: fonction indicatrice de l'interval réel non-négatif
- $0 < p \leq 1$

Problème du sac-à-dos avec contrainte en espérance (ECKP)

$$\max_{x \in \{0,1\}^n} \quad \mathbb{E}[\sum_{i=1}^{n} r_i \chi_i x_i]$$

$$\text{s.c.} \quad \mathbb{E}[\mathbb{1}_{\mathbb{R}^+}(c - \sum_{i=1}^{n} \chi_i x_i)] \geq p$$

- $\mathbb{1}_{\mathbb{R}^+}$: fonction indicatrice de l'interval réel non-négatif
- $0 < p \leq 1$

Problème du sac-à-dos avec contrainte en espérance (ECKP)

$$\max_{x\in\{0,1\}^n} \quad \mathbb{E}[\sum_{i=1}^{n} r_i \chi_i x_i]$$

$$\text{s.c.} \quad \mathbb{E}[\mathbb{1}_{\mathbb{R}^+}(c - \sum_{i=1}^{n} \chi_i x_i)] \geq p$$

- $\mathbb{1}_{\mathbb{R}^+}$: fonction indicatrice de l'interval réel non-négatif
- $0 < p \leq 1$

Problème du sac-à-dos avec contrainte en probailité (CCKP)

Problème du sac-à-dos avec contrainte en espérance (ECKP)

$$\max_{x\in\{0,1\}^n} \quad \mathbb{E}[\sum_{i=1}^{n} r_i \chi_i x_i]$$

$$\text{s.c.} \quad \mathbb{E}[\mathbb{1}_{\mathbb{R}^+}(c - \sum_{i=1}^{n} \chi_i x_i)] \geq p$$

- $\mathbb{1}_{\mathbb{R}^+}$: fonction indicatrice de l'interval réel non-négatif
- $0 < p \leq 1$

Problème du sac-à-dos avec contrainte en probailité (CCKP)

$$\mathbb{E}[\mathbb{1}_{\mathbb{R}^+}(c - \sum_{i=1}^{n} \chi_i x_i)] = \mathbb{P}\{\sum_{i=1}^{n} \chi_i x_i \leq c\}$$

Comparaison

SRKP

$$\max_{x \in \{0,1\}^n} \mathbb{E}[\sum_{i=1}^{n} r_i \chi_i x_i] - d \cdot \mathbb{E}[[\sum_{i=1}^{n} \chi_i x_i - c]^+]$$

ECKP

$$\max_{x \in \{0,1\}^n} \quad \mathbb{E}[\sum_{i=1}^{n} r_i \chi_i x_i]$$

$$\text{s.c.} \quad \mathbb{E}[\mathbb{1}_{\mathbb{R}^+}(c - \sum_{i=1}^{n} \chi_i x_i)] \geq p$$

Outline

Applications

Logistique

Applications

Logistique

- Entreprise de logistique avec capacité(s) fixe(s)

Applications

Logistique

- Entreprise de logistique avec capacité(s) fixe(s)
- Taille des objects à transporter pas connue à l'avance

Applications

Logistique

- Entreprise de logistique avec capacité(s) fixe(s)
- Taille des objects à transporter pas connue à l'avance
- $\Rightarrow$ coûts supplémentaires dans le cas de surcharge

Applications

Logistique

- Entreprise de logistique avec capacité(s) fixe(s)
- Taille des objects à transporter pas connue à l'avance
- $\Rightarrow$ coûts supplémentaires dans le cas de surcharge

Scheduling

Applications

Logistique

- Entreprise de logistique avec capacité(s) fixe(s)
- Taille des objects à transporter pas connue à l'avance
- $\Rightarrow$ coûts supplémentaires dans le cas de surcharge

Scheduling

- Ensemble de tâches possibles à traiter

Applications

Logistique

- Entreprise de logistique avec capacité(s) fixe(s)
- Taille des objects à transporter pas connue à l'avance
- $\Rightarrow$ coûts supplémentaires dans le cas de surcharge

Scheduling

- Ensemble de tâches possibles à traiter
- Chaque tâche rapporte un gain spécifique

Applications

Logistique

- Entreprise de logistique avec capacité(s) fixe(s)
- Taille des objects à transporter pas connue à l'avance
- $\Rightarrow$ coûts supplémentaires dans le cas de surcharge

Scheduling

- Ensemble de tâches possibles à traiter
- Chaque tâche rapporte un gain spécifique
- Temps d'exécution de chaque tâche pas connu à l'avance

Applications

Logistique

- Entreprise de logistique avec capacité(s) fixe(s)
- Taille des objects à transporter pas connue à l'avance
- $\Rightarrow$ coûts supplémentaires dans le cas de surcharge

Scheduling

- Ensemble de tâches possibles à traiter
- Chaque tâche rapporte un gain spécifique
- Temps d'exécution de chaque tâche pas connu à l'avance
- $\Rightarrow$ Restriction du pourcentage des cas où une limite de temps donnée est dépassée

Méthode de résolution - Schéma

Méthode de résolution - Schéma

- Algorithme "branch-and-bound"

Méthode de résolution - Schéma

- Algorithme "branch-and-bound"
- Resolution du problème relaxé pour fournir des bornes supérieurs

Méthode de résolution - Schéma

- Algorithme "branch-and-bound"
- Resolution du problème relaxé pour fournir des bornes supérieurs
- Utilisation d'un algorithme de type gradient stochastique pour résoudre la relaxation

Méthode de résolution - Schéma

- Algorithme "branch-and-bound"
- Resolution du problème relaxé pour fournir des bornes supérieurs
- Utilisation d'un algorithme de type gradient stochastique pour résoudre la relaxation
- Appliquer la méthode "Approximation par convolution" ou "Intégration par parties" pour estimer le gradient de la fonction objectif ou de la fonction de contrainte

Outline

Definition

Plobléme relaxé du sac-à-dos stochastique (avec recours simple) :

$$\max_{x \in [0,1]^n} \mathbb{E}[\sum_{i=1}^{n} r_i \chi_i x_i] - d \cdot \mathbb{E}[[\sum_{i=1}^{n} \chi_i x_i - c]^+]$$

Definition

Ploblème relaxé du sac-à-dos stochastique (avec recours simple) :

$$\max_{x \in [0,1]^n} \mathbb{E}[\sum_{i=1}^{n} r_i \chi_i x_i] - d \cdot \mathbb{E}[[\sum_{i=1}^{n} \chi_i x_i - c]^+]$$

L'algorithme du type gradient stochastique

L'algorithme gradient stochastique

L'algorithme du type gradient stochastique

L'algorithme gradient stochastique

Choisir x^0 dans X_{ad}

L'algorithme du type gradient stochastique

L'algorithme gradient stochastique

Choisir x^0 dans $X_{ad} = [0, 1]^n$

L'algorithme du type gradient stochastique

L'algorithme gradient stochastique

Choisir x^0 dans $X_{ad} = [0,1]^n$
A l'itération $k+1$, tirer $\chi^k = (\chi_1^k, ..., \chi_n^k)$ suivant la loi de χ

L'algorithme du type gradient stochastique

L'algorithme gradient stochastique

Choisir x^0 dans $X_{ad} = [0,1]^n$
A l'itération $k+1$, tirer $\chi^k = (\chi_1^k, ..., \chi_n^k)$ suivant la loi de χ
Mettre à jour x^k :

L'algorithme du type gradient stochastique

L'algorithme gradient stochastique

Choisir x^0 dans $X_{ad} = [0,1]^n$
A l'itération $k+1$, tirer $\chi^k = (\chi_1^k, ..., \chi_n^k)$ suivant la loi de χ
Mettre à jour x^k :

$$x^{k+1} = x^k + \epsilon^k r^k$$

où $r^k = \nabla j(x^k, \chi^k)$ et $(\epsilon^k)_{k \in \mathbb{N}}$ est une σ-suite

où $j(x, \chi) = \sum_j r_j \chi_j x_j - d \cdot [\sum_{i=1}^n \chi_i x_i - c]^+$

L'algorithme du type gradient stochastique

L'algorithme gradient stochastique

Choisir x^0 dans $X_{ad} = [0,1]^n$
A l'itération $k+1$, tirer $\chi^k = (\chi_1^k, ..., \chi_n^k)$ suivant la loi de χ
Mettre à jour x^k :

$$x^{k+1} = x^k + \epsilon^k r^k$$

où $r^k = \nabla j(x^k, \chi^k)$ et $(\epsilon^k)_{k \in \mathbb{N}}$ est une σ-suite
Pour tous les $i = 1, ..., n$:
Si $x_i^{k+1} > 1$ poser $x_i^{k+1} = 1$ et si $x_i^{k+1} < 0$ poser $x_i^{k+1} = 0$

où $j(x, \chi) = \sum_j r_j \chi_j x_j - d \cdot [\sum_{i=1}^n \chi_i x_i - c]^+$

L'algorithme du type Arrow-Hurwicz

L'algorithme Arrow-Hurwicz

L'algorithme du type Arrow-Hurwicz

L'algorithme Arrow-Hurwicz

Choisir $x^0 \in X_{ad} = [0,1]^n$ et $\lambda^0 \in [0, \infty)$
A l'itération $k+1$, tirer $\chi^{k+1} = (\chi_1^k, \dots, \chi_n^k)$ suivant la loi de χ

Mettre à jour x^{k+1} et λ^{k+1} :

$$x^{k+1} = x^k + \epsilon^k (r^k + (\lambda^k)^T \theta^k)$$

$$\lambda^{k+1} = \lambda^k - \rho^k (\Theta(x^{k+1}, \chi^{k+1}) - p)$$

$r^k = \nabla j(x^k, \chi^{k+1})$, $\theta^k = \nabla \Theta(x^k, \chi^{k+1})$
$(\epsilon^k)_{k \in \mathbb{N}}$ et $(\rho^k)_{k \in \mathbb{N}}$ sont des σ-suites

Pour tous les $i = 1, \dots, n$:

Si $x_i^{k+1} > 1$ poser $x_i^{k+1} = 1$ et si $x_i^{k+1} < 0$ poser $x_i^{k+1} = 0$
Si $\lambda_i^{k+1} < 0$ poser $\lambda_i^{k+1} = 0$

where $j(x, \chi) = \sum_j r_j \chi_j x_j$ and $\Theta(x, \chi) = \mathbb{1}_{\mathbb{R}^+}(c - \sum_{i=1}^n \chi_i x_i)$

L'algorithme du type Arrow-Hurwicz

L'algorithme Arrow-Hurwicz

Choisir $x^0 \in X_{ad} = [0,1]^n$ et $\lambda^0 \in [0,\infty)$
A l'itération $k+1$, tirer $\chi^{k+1} = (\chi_1^k, \ldots, \chi_n^k)$ suivant la loi de χ

Mettre à jour x^{k+1} et λ^{k+1} :

$$x^{k+1} = x^k + \epsilon^k(r^k + (\lambda^k)^T\theta^k)$$

$$\lambda^{k+1} = \lambda^k - \rho^k(\Theta(x^{k+1}, \chi^{k+1}) - p)$$

$r^k = \nabla j(x^k, \chi^{k+1})$, $\theta^k = \nabla\Theta(x^k, \chi^{k+1})$
$(\epsilon^k)_{k\in\mathbb{N}}$ et $(\rho^k)_{k\in\mathbb{N}}$ sont des σ-suites

Pour tous les $i = 1, \ldots, n$:

Si $x_i^{k+1} > 1$ poser $x_i^{k+1} = 1$ et si $x_i^{k+1} < 0$ poser $x_i^{k+1} = 0$
Si $\lambda_i^{k+1} < 0$ poser $\lambda_i^{k+1} = 0$

where $j(x,\chi) = \sum_j r_j \chi_j x_j$ and $\Theta(x,\chi) = \mathbb{1}_{\mathbb{R}^+}(c - \sum_{i=1}^n \chi_i x_i)$

Outline

Definition

Le *produit de convolution* de deux fonctions est défini comme :

Definition

Le *produit de convolution* de deux fonctions est défini comme :

$$(f * h)(x) := \int_{-\infty}^{\infty} f(y)h(x-y)\,\mathrm{d}y$$

Approximation par convolution

Approximation par convolution

Soit h une fonction paire, continue et non-négative telle que

- $\int\limits_{-\infty}^{\infty} h(x)\,\mathrm{d}x = 1$
- $\arg\max h(x) = 0$

Approximation par convolution

Soit h une fonction paire, continue et non-négative telle que

- $\int_{-\infty}^{\infty} h(x)\,\mathrm{d}x = 1$
- $\arg\max h(x) = 0$

$$h_r(x) := \frac{1}{r} h\left(\frac{x}{r}\right)$$

Approximation par convolution

Soit h une fonction paire, continue et non-négative telle que

- $\int_{-\infty}^{\infty} h(x)\,\mathrm{d}x = 1$
- $\arg\max h(x) = 0$

$$h_r(x) := \frac{1}{r} h\left(\frac{x}{r}\right)$$

$\Rightarrow$ l'approximation d'une fonction réelle et localement intégrable ρ :

$$\rho_r(x) := (\rho * h_r)(x) = \frac{1}{r} \int_{-\infty}^{\infty} \rho(y) h\left(\frac{y-x}{r}\right) \mathrm{d}y$$

Approximation par convolution

Soit h une fonction paire, continue et non-négative telle que

- $\int_{-\infty}^{\infty} h(x)\,\mathrm{d}x = 1$
- $\arg\max h(x) = 0$

$$h_r(x) := \frac{1}{r} h\left(\frac{x}{r}\right)$$

$\Rightarrow$ l'approximation d'une fonction réelle et localement intégrable ρ :

$$\rho_r(x) := (\rho * h_r)(x) = \frac{1}{r} \int_{-\infty}^{\infty} \rho(y) h\left(\frac{y-x}{r}\right)\,\mathrm{d}y$$

Quand $\rho = \mathbb{1}_{\mathbb{R}^+}$:

$$(\rho_r)'(x) = \frac{1}{r^2} \int_0^{\infty} h'\left(\frac{y-x}{r}\right)\,\mathrm{d}y = -\frac{1}{r} h\left(\frac{-x}{r}\right) = -\frac{1}{r} h\left(\frac{x}{r}\right)$$

$$h(x) := \frac{3}{4}(1 - x^2)\mathbb{1}_1(x)$$

($\mathbb{1}_1$: fonction indicatrice de l'interval $]-1, 1[$)

$$h(x) := \frac{3}{4}(1 - x^2)\mathbb{1}_1(x)$$

($\mathbb{1}_1$: fonction indicatrice de l'interval $]-1, 1[$)

Gradient approximé de la fonction objectif du SRKP

$$\nabla_x j(x, \chi) = (r_1\chi_1, \ldots, r_n\chi_n)^T +$$

$$d \cdot \left(\frac{3}{4r} \left(1 - \left(\frac{g(x,\chi)}{r} \right)^2 \right) \mathbb{1}_1\left(\frac{g(x,\chi)}{r} \right) \chi \cdot g(x,\chi) - \mathbb{1}_{\mathbb{R}^+}(g(x,\chi)) \cdot \chi \right)$$

$$h(x) := \frac{3}{4}(1-x^2)\mathbb{1}_1(x)$$

($\mathbb{1}_1$: fonction indicatrice de l'interval $]-1,1[$)

Gradient approximé de la fonction objectif du SRKP

$$\nabla_x j(x,\chi) = (r_1\chi_1, \ldots, r_n\chi_n)^T +$$

$$d \cdot \left(\frac{3}{4r} \left(1 - \left(\frac{g(x,\chi)}{r} \right)^2 \right) \mathbb{1}_1\left(\frac{g(x,\chi)}{r} \right) \chi \cdot g(x,\chi) - \mathbb{1}_{\mathbb{R}^+}(g(x,\chi)) \cdot \chi \right)$$

where $g(x,\chi) := \sum_{i=1}^n \chi_i x_i - c$

Outline

Intégration par parties - Idée sur l'exemple du SRKP

Intégration par parties - Idée sur l'exemple du SRKP

Intégration par parties

- Utiliser Intégration par parties pour reformuler $\mathbb{E}[[\chi_i x_i - c]^+]$

Intégration par parties - Idée sur l'exemple du SRKP

Intégration par parties

- Utiliser Intégration par parties pour reformuler $\mathbb{E}[[\chi_i x_i - c]^+] = \mathbb{E}[\tilde{j}(x, \chi)]$

Intégration par parties - Idée sur l'exemple du SRKP

Intégration par parties

- Utiliser Intégration par parties pour reformuler $\mathbb{E}[[\chi_i x_i - c]^+] = \mathbb{E}[\tilde{j}(x,\chi)]$
- $\tilde{j}(x,\chi) = \mathbb{Y}_{\mathbb{R}^+}(\chi_i x_i - c) h(x,\chi)$

Intégration par parties - Idée sur l'exemple du SRKP

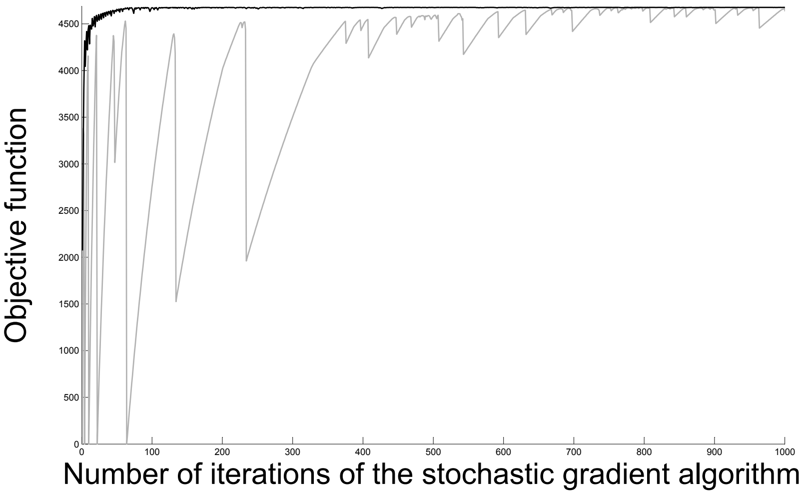
Intégration par parties

- Utiliser Intégration par parties pour reformuler $\mathbb{E}[[\chi_i x_i - c]^+] = \mathbb{E}[\tilde{j}(x,\chi)]$
- $\tilde{j}(x,\chi) = \mathbb{Y}_{\mathbb{R}^+}(\chi_i x_i - c) h(x,\chi)$
- Calculer gradient de $\tilde{j}(x,\chi)$

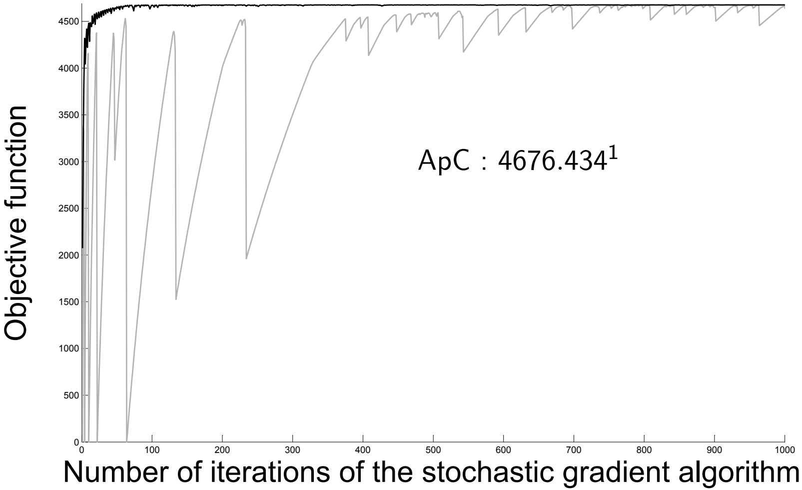
Intégration par parties - Idée sur l'exemple du SRKP

Intégration par parties

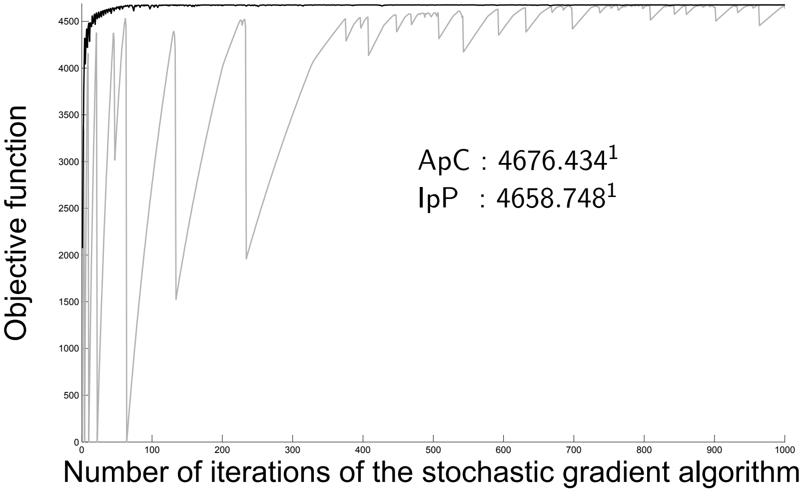
- Utiliser Intégration par parties pour reformuler $\mathbb{E}[[\chi_i x_i - c]^+] = \mathbb{E}[\tilde{j}(x,\chi)]$
- $\tilde{j}(x,\chi) = \mathbb{Y}_{\mathbb{R}^+}(\chi_i x_i - c) h(x,\chi)$
- Calculer gradient de $\tilde{j}(x,\chi)$
- Utiliser $\nabla_x \tilde{j}(x,\chi)$ dans l'algorithme au lieu de $\nabla_x j(x,\chi)$


Objective function
4500
4000
3500
3000
2500
2000
1500
1000
500
0
0
100
200
300
400
500
600
700
800
900
1000
Number of iterations of the stochastic gradient algorithm

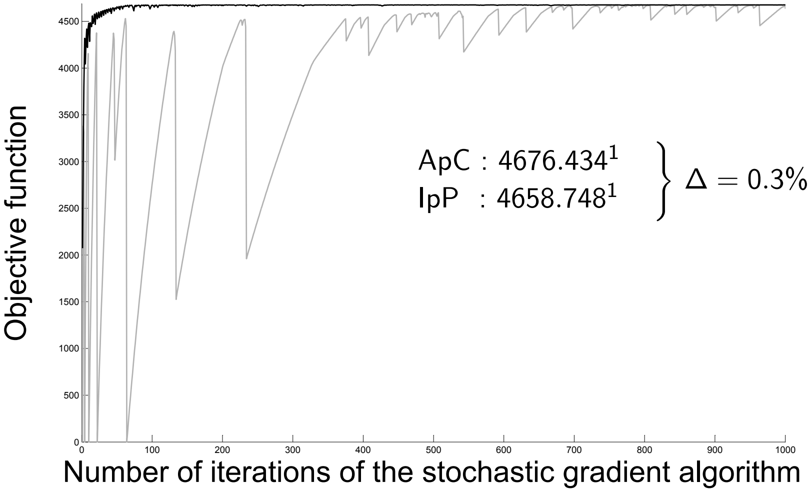
1



[1] moyenne sur 1000 tests

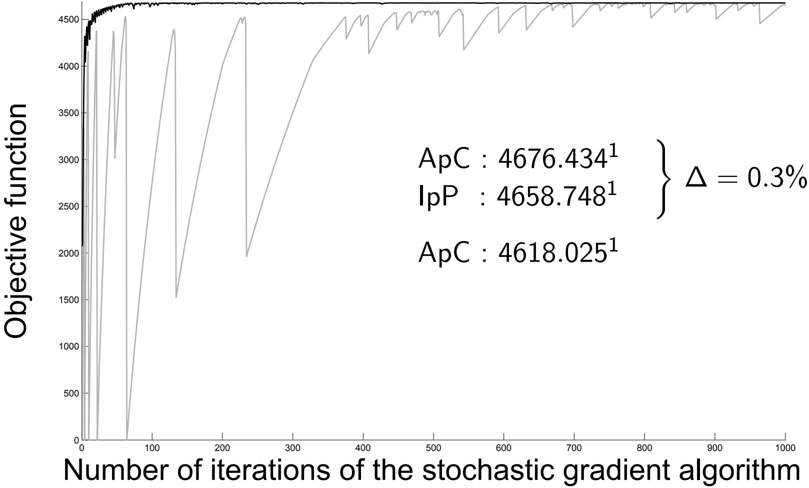


[1] moyenne sur 1000 tests



$$\left.\begin{array}{ll}\text{ApC} & : 4676.434^{1} \\ \text{IpP} & : 4658.748^{1}\end{array}\right\} \Delta = 0.3\%$$

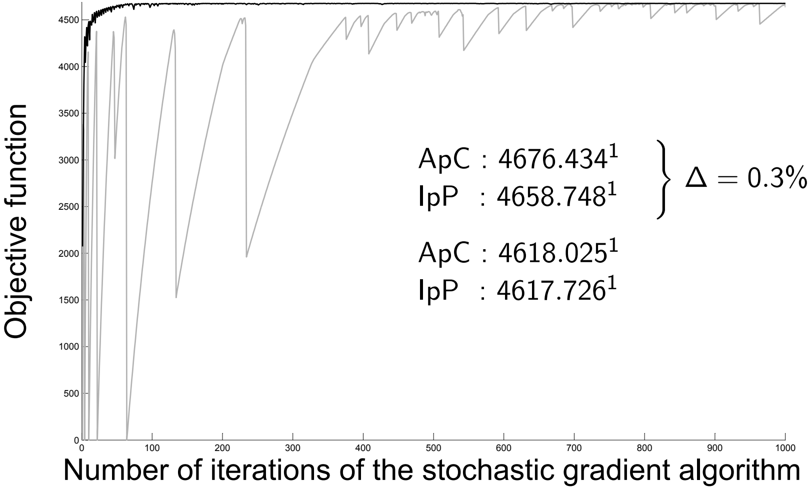
[1] moyenne sur 1000 tests



$$\left.\begin{array}{ll} \text{ApC} & : 4676.434^{1} \\ \text{IpP} & : 4658.748^{1} \end{array}\right\} \Delta = 0.3\%$$

$$\text{ApC} : 4618.025^{1}$$

[1] moyenne sur 1000 tests



$$\left.\begin{array}{ll}\text{ApC} & : 4676.434^{1} \\ \text{IpP} & : 4658.748^{1}\end{array}\right\} \Delta = 0.3\%$$

ApC : 4618.025[1]

IpP : 4617.726[1]

[1] moyenne sur 1000 tests

Amélioration de la convergence de IpP

Amélioration de la convergence de IpP

Gradient stochastique à pas fix

Amélioration de la convergence de IpP

Gradient stochastique à pas fix

Mettre à jour x^k :

$$x^{k+1} = x^k + \epsilon^k r^k$$

où $r^k = \nabla_x \tilde{j}(x^k, \chi^k)$

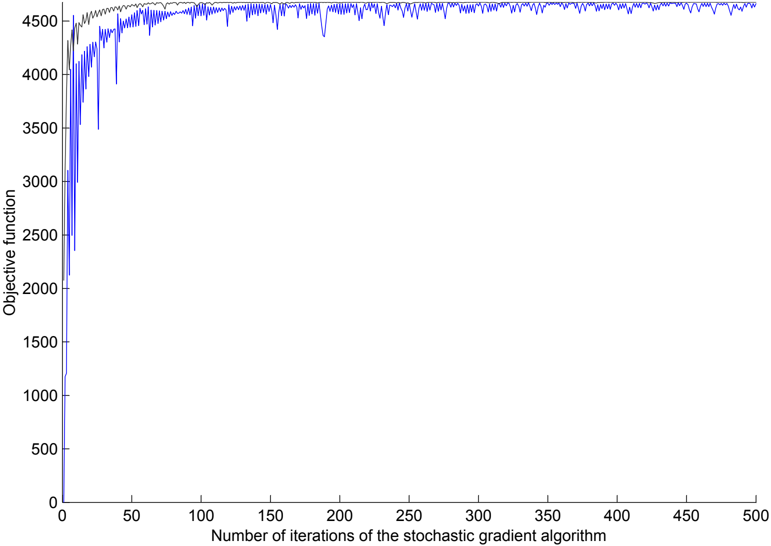
Amélioration de la convergence de IpP

Gradient stochastique à pas variable

Mettre à jour x^k :

$$x^{k+1} = x^k + \epsilon^k \frac{r^k}{\|r^k\|}$$

où $r^k = \nabla_x \tilde{j}(x^k, \chi^k)$


4500
4000
3500
3000
2500
2000
1500
1000
500
0
Objective function
0
50
100
150
200
250
300
350
400
450
500
Number of iterations of the stochastic gradient algorithm

Merci !