

Introduction aux grands systèmes et rappels

Plan du cours

- 1 **Présentation des grands systèmes**
 - Grands systèmes en optimisation
 - Problèmes prototypes
- 2 **Rappels d'optimisation convexe**
 - Propriétés élémentaires
 - Optimisation sans contrainte explicite
 - Optimisation avec contraintes explicites
- 3 **Séance de travaux dirigés**
 - Stabilité du Lagrangien
 - Interprétation géométrique de l'algorithme d'Uzawa
 - Optimisation d'un grand réseau de distribution d'eau

- 1 Présentation des grands systèmes
 - Grands systèmes en optimisation
 - Problèmes prototypes
- 2 Rappels d'optimisation convexe
 - Propriétés élémentaires
 - Optimisation sans contrainte explicite
 - Optimisation avec contraintes explicites
- 3 Séance de travaux dirigés
 - Stabilité du Lagrangien
 - Interprétation géométrique de l'algorithme d'Uzawa
 - Optimisation d'un grand réseau de distribution d'eau

- 1 Présentation des grands systèmes
 - Grands systèmes en optimisation
 - Problèmes prototypes
- 2 Rappels d'optimisation convexe
 - Propriétés élémentaires
 - Optimisation sans contrainte explicite
 - Optimisation avec contraintes explicites
- 3 Séance de travaux dirigés
 - Stabilité du Lagrangien
 - Interprétation géométrique de l'algorithme d'Uzawa
 - Optimisation d'un grand réseau de distribution d'eau

Qu'est-ce qu'un grand système ?

Du point de vue de l'**optimisation**, un **grand système** est tel que :

- il comporte un grand nombre de variables et de contraintes,
- il présente une hétérogénéité spatiale et/ou temporelle,
- le nombre de décideurs intervenant sur le système est grand.

Impossibilité d'utiliser les techniques classiques de l'optimisation.

Cette impossibilité est d'ordre **methodologique**, liée aux ressources (CPU, mémoire) nécessaires pour résoudre le problème.

- **Programmation dynamique**

Exemple de la gestion optimale d'une vallée hydraulique

Barrages	2	3	4	5
CPU	≈ 1 minute	≈ 1 heure	≈ 1 jour	≈ 1 an

- **Conditions de Karush-Kuhn-Tucker**

Nombre d'alternatives avec N contraintes inégalité : 2^N .

Qu'est-ce qu'un grand système ?

Du point de vue de l'**optimisation**, un **grand système** est tel que :

- il comporte un grand nombre de variables et de contraintes,
- il présente une hétérogénéité spatiale et/ou temporelle,
- le nombre de décideurs intervenant sur le système est grand.

Impossibilité d'utiliser les techniques classiques de l'optimisation.

Cette impossibilité est d'ordre **méthodologique**, liée aux ressources (CPU, mémoire) nécessaires pour résoudre le problème.

- **Programmation dynamique**

Exemple de la gestion optimale d'une vallée hydraulique

Barrages	2	3	4	5
CPU	≈ 1 minute	≈ 1 heure	≈ 1 jour	≈ 1 an

- **Conditions de Karush-Kuhn-Tucker**

Nombre d'alternatives avec N contraintes inégalité : 2^N .

Pourquoi et comment optimiser un grand système ?

En général, les grands systèmes sont tels que :

- ils sont **couplés** et ne peuvent être pilotés localement,
- ils représentent des **enjeux économiques importants**.

On se limite aux systèmes tels que :

- **un seul critère** doit être minimisé,
- **un seul décideur** agit sur le système.

C'est la **situation classique** de l'optimisation déterministe.

Pour optimiser un tel système, on va le **découper** en plusieurs petits sous-systèmes :

- ① chaque sous-système local sera résolu par les techniques classiques de l'optimisation (**décomposition**),
- ② la comparaison des solutions locales servira à mettre à jour les sous-systèmes pour obtenir la solution globale (**coordination**).

La **décomposition/coordination** est donc un processus **itératif**.

Pourquoi et comment optimiser un grand système ?

En général, les grands systèmes sont tels que :

- ils sont **couplés** et ne peuvent être pilotés localement,
- ils représentent des **enjeux économiques importants**.

On se limite aux systèmes tels que :

- **un seul critère** doit être minimisé,
- **un seul décideur** agit sur le système.

C'est la **situation classique** de l'optimisation déterministe.

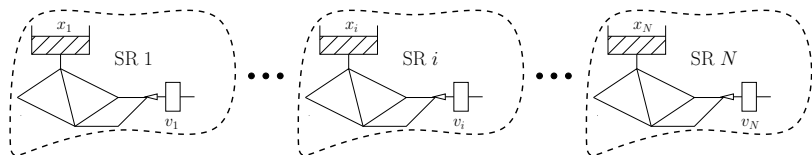
Pour optimiser un tel système, on va le **découper** en plusieurs petits sous-systèmes :

- 1 chaque sous-système local sera résolu par les techniques classiques de l'optimisation (**décomposition**),
- 2 la comparaison des solutions locales servira à mettre à jour les sous-systèmes pour obtenir la solution globale (**coordination**).

La **décomposition/coordination** est donc un processus **itératif**.

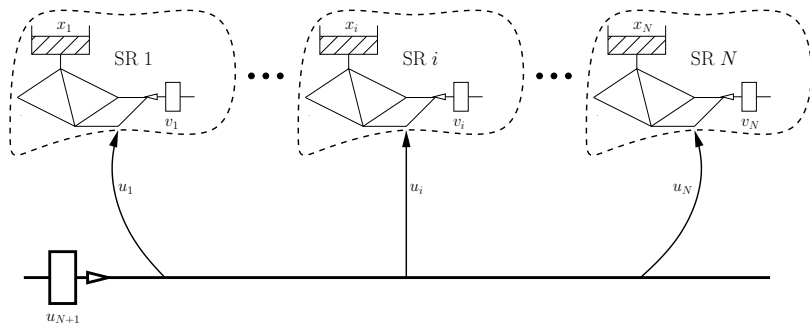
- 1 Présentation des grands systèmes
 - Grands systèmes en optimisation
 - Problèmes prototypes
- 2 Rappels d'optimisation convexe
 - Propriétés élémentaires
 - Optimisation sans contrainte explicite
 - Optimisation avec contraintes explicites
- 3 Séance de travaux dirigés
 - Stabilité du Lagrangien
 - Interprétation géométrique de l'algorithme d'Uzawa
 - Optimisation d'un grand réseau de distribution d'eau

Réseaux de distribution d'eau interconnectés



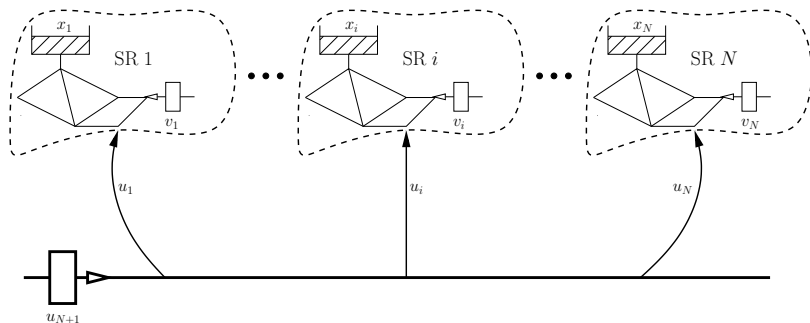
Le réseau d'eau est constitué initialement de N petits sous-réseaux **indépendants** les uns des autres, faciles à optimiser.

Réseaux de distribution d'eau interconnectés



Pour des raisons de sécurité de l'approvisionnement, on **connecte** ces sous-réseaux avec une ressource en eau supplémentaire.

Réseaux de distribution d'eau interconnectés



L'optimisation d'un grand réseau de distribution d'eau à l'échelle régionale sera étudiée durant les séances de **travaux dirigés**.

Exemple canonique de grands systèmes

Soit une entreprise comportant N unités de production **autonomes**. L'unité i est pilotée par une variable u_i , respectant les contraintes $u_i \in U_i^{\text{ad}}$. La **production** et le **coût** associés à u_i sont notés $\Theta_i(u_i)$ et $J_i(u_i)$. L'objectif de l'entreprise est de **minimiser** son **coût total** de production, sous contrainte d'une **production totale** égale à θ .

$$\min_{(u_1, \dots, u_N) \in U_1^{\text{ad}} \times \dots \times U_N^{\text{ad}}} \sum_{i=1}^N J_i(u_i),$$

sous la contrainte :

$$\sum_{i=1}^N \Theta_i(u_i) = \theta.$$

Le critère et la contrainte sont **additifs** en i . C'est une situation très favorable pour les méthodes de **décomposition/coordination**.

Exemple canonique de grands systèmes

Soit une entreprise comportant N unités de production **autonomes**. L'unité i est pilotée par une variable u_i , respectant les contraintes $u_i \in U_i^{\text{ad}}$. La **production** et le **coût** associés à u_i sont notés $\Theta_i(u_i)$ et $J_i(u_i)$. L'objectif de l'entreprise est de **minimiser** son **coût total** de production, sous contrainte d'une **production totale** égale à θ .

$$\min_{(u_1, \dots, u_N) \in U_1^{\text{ad}} \times \dots \times U_N^{\text{ad}}} \sum_{i=1}^N J_i(u_i),$$

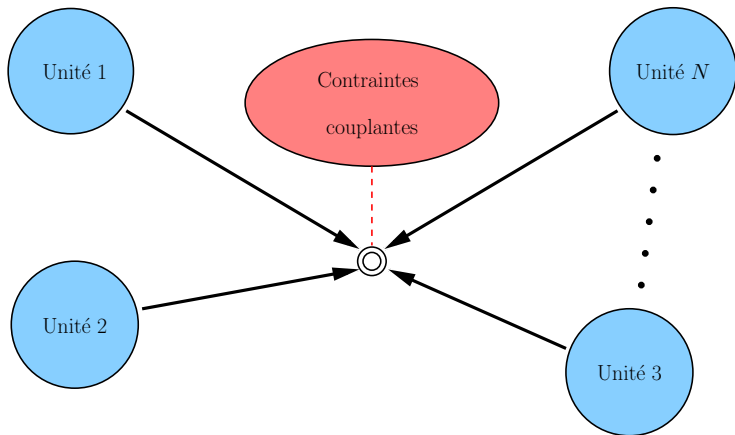
sous la contrainte :

$$\sum_{i=1}^N \Theta_i(u_i) = \theta.$$

Le critère et la contrainte sont **additifs** en i . C'est une situation très **favorable** pour les méthodes de **décomposition/coordination**.

Exemple canonique de grands systèmes

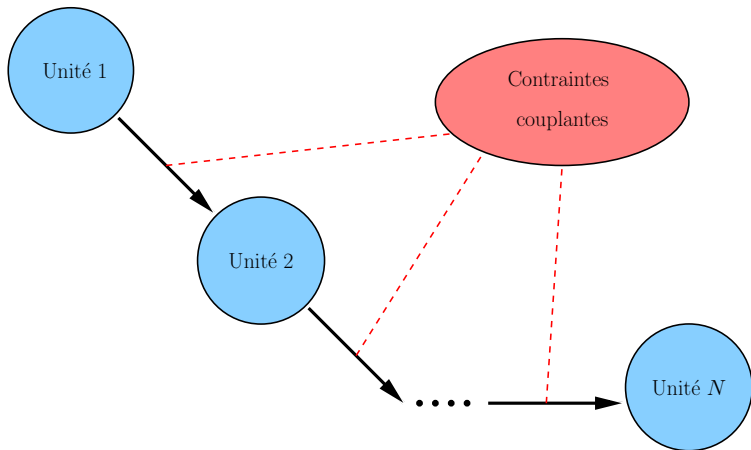
II



Systeme en étoile

Exemple canonique de grands systèmes

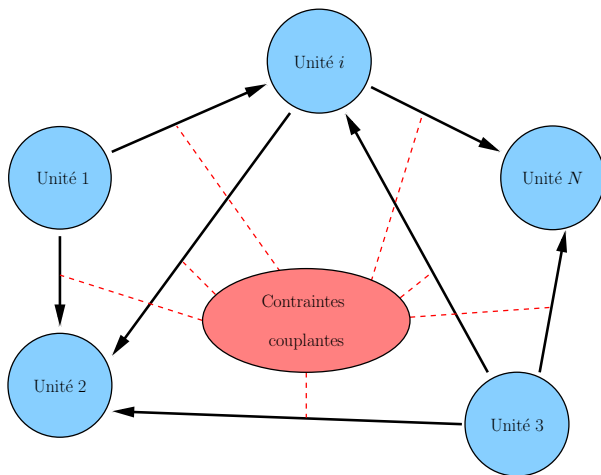
III



Système en cascade

Exemple canonique de grands systèmes

IV



Système en réseau

Contenu du cours

- **Séance No.1**

Rappels d'optimisation convexe et de dualité.

- **Séance No.2 & 3**

Approche « intuitive » de la décomposition-coordination :
le cas additif.

- **Séance No.4 & 5**

Approche mathématique de la décomposition-coordination :
le cas général.

- **Séance No.6**

Évaluation des connaissances.

- 1 Présentation des grands systèmes
 - Grands systèmes en optimisation
 - Problèmes prototypes
- 2 Rappels d'optimisation convexe
 - Propriétés élémentaires
 - Optimisation sans contrainte explicite
 - Optimisation avec contraintes explicites
- 3 Séance de travaux dirigés
 - Stabilité du Lagrangien
 - Interprétation géométrique de l'algorithme d'Uzawa
 - Optimisation d'un grand réseau de distribution d'eau

Deux problèmes d'optimisation

On va rappeler les principaux résultats d'optimisation associés aux deux problèmes suivants.

- Optimisation **sans contrainte explicite**

$$\min_{u \in U^{\text{ad}} \subset \mathcal{U}} J(u) . \quad (\mathcal{P}_S)$$

- Optimisation **avec contraintes explicites**

$$\min_{u \in U^{\text{ad}} \subset \mathcal{U}} J(u) \quad \text{sous} \quad \Theta(u) \in -C \subset \mathcal{V} . \quad (\mathcal{P}_A)$$

La notation $\Theta(u) \in -C$ englobe le cas des contraintes **égalité** ($C = \{0\}$), des contraintes **inégalité** en dimension finie $\mathcal{V} = \mathbb{R}^m$ ($C = \mathbb{R}_+^m$), et permet de les généraliser en **dimension infinie**.

- 1 Présentation des grands systèmes
 - Grands systèmes en optimisation
 - Problèmes prototypes
- 2 Rappels d'optimisation convexe
 - Propriétés élémentaires
 - Optimisation sans contrainte explicite
 - Optimisation avec contraintes explicites
- 3 Séance de travaux dirigés
 - Stabilité du Lagrangien
 - Interprétation géométrique de l'algorithme d'Uzawa
 - Optimisation d'un grand réseau de distribution d'eau

Espaces et ensembles

- \mathcal{U} et \mathcal{V} sont des espaces de **Hilbert**.
 - Exemple en dimension finie : \mathbb{R}^n .
 - Exemple en dimension infinie : $L^2([0, T], \mathbb{R}^n)$.
- $U^{\text{ad}} \subset \mathcal{U}$ est un ensemble **convexe**, **fermé** et non vide.
- $C \subset \mathcal{V}$ est un **cône** ($\forall \alpha \geq 0, v \in C \Rightarrow \alpha v \in C$) que l'on suppose **convexe**, **fermé** et vérifiant : $C \cap (-C) = \{0\}$.
 - Le **cône dual** de C est : $C^* = \{p \in \mathcal{V}, \langle p, v \rangle \geq 0, \forall v \in C\}$.
Dans le cas $\mathcal{V} = \mathbb{R}^m$, le cône dual associé
 - à des contraintes **égalité** ($C = \{0\}$) est : $C^* = \mathbb{R}^m$,
 - à des contraintes **inégalité** ($C = \mathbb{R}_+^m$) est : $C^* = \mathbb{R}_+^m$.

Espaces et ensembles

- \mathcal{U} et \mathcal{V} sont des espaces de **Hilbert**.
 - Exemple en dimension finie : \mathbb{R}^n .
 - Exemple en dimension infinie : $L^2([0, T], \mathbb{R}^n)$.
- $\mathcal{U}^{\text{ad}} \subset \mathcal{U}$ est un ensemble **convexe**, **fermé** et non vide.
- $C \subset \mathcal{V}$ est un **cône** ($\forall \alpha \geq 0, v \in C \Rightarrow \alpha v \in C$) que l'on suppose **convexe**, **fermé** et vérifiant : $C \cap (-C) = \{0\}$.
- Le **cône dual** de C est : $C^* = \{p \in \mathcal{V}, \langle p, v \rangle \geq 0, \forall v \in C\}$.
Dans le cas $\mathcal{V} = \mathbb{R}^m$, le cône dual associé
 - à des contraintes **égalité** ($C = \{0\}$) est : $C^* = \mathbb{R}^m$,
 - à des contraintes **inégalité** ($C = \mathbb{R}_+^m$) est : $C^* = \mathbb{R}_+^m$.

Critère

Le critère à optimiser est une fonction $J : \mathcal{U} \rightarrow \overline{\mathbb{R}}$. On appelle **domaine** de J et l'on note $\text{dom}J$ l'ensemble :

$$\text{dom}J = \{u \in \mathcal{U}, J(u) < +\infty\} .$$

On suppose que J est une fonction **propre** (non identiquement égale à $+\infty$ et ne prenant pas la valeur $-\infty$) et que minimiser la fonction J sur l'ensemble U^{ad} a un sens :

$$\text{dom}J \cap U^{\text{ad}} \neq \emptyset .$$

On fait alors sur la fonction J des hypothèses de convexité, de continuité, de différentiabilité et de comportement à l'infini.

Critère

Le critère à optimiser est une fonction $J : \mathcal{U} \rightarrow \overline{\mathbb{R}}$. On appelle **domaine** de J et l'on note $\text{dom}J$ l'ensemble :

$$\text{dom}J = \{u \in \mathcal{U}, J(u) < +\infty\} .$$

On suppose que J est une fonction **propre** (non identiquement égale à $+\infty$ et ne prenant pas la valeur $-\infty$) et que minimiser la fonction J sur l'ensemble U^{ad} a un sens :

$$\text{dom}J \cap U^{\text{ad}} \neq \emptyset .$$

On fait alors sur la fonction J des hypothèses de **convexité**, de **continuité**, de **différentiabilité** et de **comportement à l'infini**.

Critère



- Hypothèse de **convexité**

$$J(\alpha u_1 + (1 - \alpha)u_2) \leq \alpha J(u_1) + (1 - \alpha)J(u_2) \quad \forall \alpha \in [0, 1],$$

voire, de **stricte** convexité ou de **forte** convexité.

- Hypothèse de **semi-continuité inférieure (s.c.i.)**

$$\liminf_{u \rightarrow u_0} J(u) \geq J(u_0),$$

voire, de **continuité** ou de **Lipschitz**.

- Hypothèse de **Gâteaux-différentiabilité**

$$\lim_{\epsilon \rightarrow 0_+} \frac{J(u_0 + \epsilon d) - J(u_0)}{\epsilon} = \langle \nabla J(u_0), d \rangle \quad \forall d \in \mathcal{U},$$

voire, de **sous-différentiabilité** ou de **Fréchet-différentiabilité**.

- Hypothèse de **coercivité** sur U^{ad}

$$\lim_{\|u\| \rightarrow +\infty, u \in U^{\text{ad}}} J(u) = +\infty.$$

Contraintes

Les contraintes sont **modélisées** par une application $\Theta : \mathcal{U} \rightarrow \mathcal{V}$, vérifiant des **hypothèses** de même nature que J , mais **adaptées** au fait que Θ prend ses valeurs dans l'opposé du cône $C \subset \mathcal{V}$.

- Hypothèse de **C-convexité**

$$\Theta(\alpha u_1 + (1-\alpha)u_2) - \alpha\Theta(u_1) - (1-\alpha)\Theta(u_2) \in -C \quad \forall \alpha \in [0, 1].$$

- Hypothèse de **continuité**

$$\lim_{u \rightarrow u_0} \Theta(u) = \Theta(u_0).$$

- Hypothèse de **Gâteaux-différentiabilité**

$$\lim_{\epsilon \rightarrow 0_+} \frac{\Theta(u_0 + \epsilon d) - \Theta(u_0)}{\epsilon} = \Theta'(u_0).d \quad \forall d \in \mathcal{U}.$$

Remarque. **C-convexité** $\iff u \mapsto \langle p, \Theta(u) \rangle$ **convexe** pour tout $p \in C^*$.

- 1 Présentation des grands systèmes
 - Grands systèmes en optimisation
 - Problèmes prototypes
- 2 Rappels d'optimisation convexe
 - Propriétés élémentaires
 - **Optimisation sans contrainte explicite**
 - Optimisation avec contraintes explicites
- 3 Séance de travaux dirigés
 - Stabilité du Lagrangien
 - Interprétation géométrique de l'algorithme d'Uzawa
 - Optimisation d'un grand réseau de distribution d'eau

Existence et unicité, caractérisation et calcul

On donne une réponse aux questions de l'**existence**, de l'**unicité**, de la **caractérisation** et du **calcul** de la solution du problème (\mathcal{P}_S).

$$\min_{u \in U^{\text{ad}} \subset \mathcal{U}} J(u) .$$

- **Existence.** En dimension finie, supposer que la fonction J est s.c.i., que l'ensemble U^{ad} est fermé et que la fonction J est coercive sur U^{ad} permet de garantir l'existence d'une solution du problème.¹
- **Unicité.** Supposer en plus des hypothèses précédentes que la fonction J est **strictement convexe** et que l'ensemble U^{ad} est convexe garantit l'unicité de la solution du problème.

¹ En dimension infinie, ces conditions ne sont pas suffisantes, et il faut ajouter des hypothèses de convexité.

Existence et unicité, caractérisation et calcul

On donne une réponse aux questions de l'**existence**, de l'**unicité**, de la **caractérisation** et du **calcul** de la solution du problème (\mathcal{P}_S).

$$\min_{u \in U^{\text{ad}} \subset \mathcal{U}} J(u) .$$

- **Existence.** En **dimension finie**, supposer que la fonction J est **s.c.i.**, que l'ensemble U^{ad} est **fermé** et que la fonction J est **coercive** sur U^{ad} permet de garantir l'existence d'une solution du problème.¹
- **Unicité.** Supposer en plus des hypothèses précédentes que la fonction J est **strictement convexe** et que l'ensemble U^{ad} est **convexe** garantit l'unicité de la solution du problème.

1. En **dimension infinie**, ces conditions ne sont pas suffisantes, et il faut ajouter des hypothèses de **convexité**...

Existence et unicité, caractérisation et calcul

II

- **Caractérisation.** Si J est de plus **Gâteaux-différentiable**, une condition **nécessaire et suffisante** pour que u^\sharp soit solution est que l'**inégalité variationnelle** suivante soit vérifiée :

$$\langle \nabla J(u^\sharp), u - u^\sharp \rangle \geq 0 \quad \forall u \in U^{\text{ad}} .$$

Cas particulier. En l'**absence de toute contrainte** ($U^{\text{ad}} = \mathcal{U}$), cette condition s'écrit : $\nabla J(u^\sharp) = 0$.

- **Calcul.** Si J est **fortement convexe** (de module a) à **gradient Lipschitzien** (de constante A), l'algorithme du gradient :

$$u^{(k+1)} = \text{proj}_{U^{\text{ad}}} (u^{(k)} - \rho \nabla J(u^{(k)})) ,$$

converge vers l'unique solution u^\sharp du problème pourvu que le pas ρ de l'algorithme vérifie la condition : $\rho \in]0, \frac{2a}{A^2} [$.

Il existe bien sûr d'autres algorithmes plus performants que le gradient (**gradient conjugué, quasi-Newton, Newton...**).

- 1 Présentation des grands systèmes
 - Grands systèmes en optimisation
 - Problèmes prototypes
- 2 Rappels d'optimisation convexe
 - Propriétés élémentaires
 - Optimisation sans contrainte explicite
 - Optimisation avec contraintes explicites
- 3 Séance de travaux dirigés
 - Stabilité du Lagrangien
 - Interprétation géométrique de l'algorithme d'Uzawa
 - Optimisation d'un grand réseau de distribution d'eau

Cadre théorique général

Dans le cas du problème (\mathcal{P}_A) ,

$$\min_{u \in U^{\text{ad}} \subset \mathcal{U}} J(u) \quad \text{sous} \quad \Theta(u) \in -C \subset \mathcal{V},$$

on peut se ramener au cas **sans contrainte explicite** en posant :

$$U^\ominus = \{u \in \mathcal{U}, \Theta(u) \in -C\},$$

et en effectuant la minimisation de J sur l'ensemble $U^{\text{ad}} \cap U^\ominus$.²

Cependant, l'utilisation d'une expression analytique des contraintes permet de donner une caractérisation opératoire de la solution du problème et permet de disposer d'algorithmes efficaces.

Par ailleurs, projeter sur une intersection d'ensembles convexes correspond à un problème d'optimisation en général difficile.

2. Les hypothèses faites sur le cône C et la fonction Θ impliquent que l'ensemble U^\ominus est un ensemble convexe fermé de \mathcal{U} .

Cadre théorique général

Dans le cas du problème (\mathcal{P}_A) ,

$$\min_{u \in U^{\text{ad}} \subset \mathcal{U}} J(u) \quad \text{sous} \quad \Theta(u) \in -C \subset \mathcal{V},$$

on peut se ramener au cas **sans contrainte explicite** en posant :

$$U^\ominus = \{u \in \mathcal{U}, \Theta(u) \in -C\},$$

et en effectuant la minimisation de J sur l'ensemble $U^{\text{ad}} \cap U^\ominus$.²

Cependant, l'utilisation d'une **expression analytique** des contraintes permet de donner une **caractérisation opératoire** de la solution du problème et permet de disposer d'**algorithmes efficaces**.

Par ailleurs, projeter sur une intersection d'ensembles convexes correspond à un problème d'optimisation en général difficile...

2. Les hypothèses faites sur le cône C et la fonction Θ impliquent que l'ensemble U^\ominus est un ensemble convexe fermé de \mathcal{U} .

Qualification des contraintes

Pour **caractériser** les solutions du problème (\mathcal{P}_A) , on introduit, en plus des hypothèses déjà faites de convexité, de continuité et de différentiabilité de J et Θ , une condition dite de **qualification des contraintes**. La condition générale que l'on utilise est :

$$0 \in \text{int}(\Theta(U^{\text{ad}} \cap \text{dom}J) + C) ,$$

où **int** désigne l'**intérieur topologique** d'un ensemble.

À partir de cette condition générale, on retrouve les conditions classiques dans le cas de **contraintes de type égalité et inégalité** (on suppose pour simplifier que $\text{dom}J = U$) :

Cas égalité $C = \{0\}$: $0 \in \text{int}(\Theta(U^{\text{ad}}))$.

Cas inégalité $C = \mathbb{R}_+^m$: $\exists u_0 \in U^{\text{ad}}, \Theta(u_0) \in \text{int}(-C)$.

Qualification des contraintes

Pour **caractériser** les solutions du problème (\mathcal{P}_A) , on introduit, en plus des hypothèses déjà faites de convexité, de continuité et de différentiabilité de J et Θ , une condition dite de **qualification des contraintes**. La condition générale que l'on utilise est :

$$0 \in \text{int}(\Theta(U^{\text{ad}} \cap \text{dom}J) + C) ,$$

où **int** désigne l'**intérieur topologique** d'un ensemble.

À partir de cette condition générale, on retrouve les conditions classiques dans le cas de **contraintes de type égalité et inégalité** (on suppose pour simplifier que $\text{dom}J = \mathcal{U}$) :

Cas égalité $C = \{0\}$: $0 \in \text{int}(\Theta(U^{\text{ad}}))$.

Cas inégalité $C = \mathbb{R}_+^m$: $\exists u_0 \in U^{\text{ad}}, \Theta(u_0) \in \text{int}(-C)$.

Conditions de Karush-Kuhn-Tucker

Sous l'hypothèse de **qualification**, une solution u^\sharp de (\mathcal{P}_A) est caractérisée par l'**existence** d'un $p^\sharp \in \mathcal{V}$ appelé **multiplicateur**, vérifiant les conditions de **Karush-Kuhn-Tucker (KKT)** :

$$\begin{aligned}\langle \nabla J(u^\sharp) + (\Theta'(u^\sharp))^\top \cdot p^\sharp, u - u^\sharp \rangle &\geq 0 \quad \forall u \in U^{\text{ad}}, \\ \Theta(u^\sharp) &\in -C, \quad p^\sharp \in C^*, \\ \langle p^\sharp, \Theta(u^\sharp) \rangle &= 0.\end{aligned}$$

La dernière condition (dite condition des écarts complémentaires) est de type combinatoire pour les contraintes de type inégalité : avec $C = \mathbb{R}_+^m$, elle s'écrit :

$$\sum_{j=1}^m p_j^\sharp \Theta_j(u^\sharp) = 0,$$

avec $\Theta_j(u^\sharp) \leq 0$ et $p_j^\sharp \geq 0$. Chaque terme $p_j^\sharp \Theta_j(u^\sharp)$ de cette somme est donc nul, et la condition des écarts complémentaires est donc équivalente à :

$$p_j^\sharp = 0 \quad \text{ou} \quad \Theta_j(u^\sharp) = 0 \quad \forall j = 1, \dots, m.$$

Conditions de Karush-Kuhn-Tucker

Sous l'hypothèse de **qualification**, une solution u^\sharp de (\mathcal{P}_A) est caractérisée par l'**existence** d'un $p^\sharp \in \mathcal{V}$ appelé **multiplicateur**, vérifiant les conditions de **Karush-Kuhn-Tucker (KKT)** :

$$\begin{aligned}\langle \nabla J(u^\sharp) + (\Theta'(u^\sharp))^\top \cdot p^\sharp, u - u^\sharp \rangle &\geq 0 \quad \forall u \in U^{\text{ad}}, \\ \Theta(u^\sharp) &\in -C, \quad p^\sharp \in C^*, \\ \langle p^\sharp, \Theta(u^\sharp) \rangle &= 0.\end{aligned}$$

La dernière condition (dite condition des **écarts complémentaires**) est de type **combinatoire** pour les contraintes de type inégalité : avec $C = \mathbb{R}_+^m$, elle s'écrit :

$$\sum_{j=1}^m p_j^\sharp \Theta_j(u^\sharp) = 0,$$

avec $\Theta_j(u^\sharp) \leq 0$ et $p_j^\sharp \geq 0$. Chaque terme $p_j^\sharp \Theta_j(u^\sharp)$ de cette somme est donc nul, et la **condition des écarts complémentaires** est donc équivalente à :

$$p_j^\sharp = 0 \quad \text{ou} \quad \Theta_j(u^\sharp) = 0 \quad \forall j = 1, \dots, m.$$

Conditions de Karush-Kuhn-Tucker

II

Les **gradients partiels** par rapport à u et p du **Lagrangien** L du problème (avec $L(u, p) = J(u) + \langle p, \Theta(u) \rangle$) ont pour expression :

$$\nabla_u L(u, p) = \nabla J(u) + (\Theta'(u))^T \cdot p \quad , \quad \nabla_p L(u, p) = \Theta(u) .$$

Dans le cas $U^{\text{ad}} = \mathcal{U}$ (toutes les contraintes sont données par Θ), les conditions de KKT prennent les formes particulières suivantes :

- pour les contraintes égalité $\Theta(u) = 0$:

$$\nabla_u L(u^{\text{opt}}, p^{\text{opt}}) = 0 ,$$

$$\nabla_p L(u^{\text{opt}}, p^{\text{opt}}) = 0 ;$$

- pour les contraintes inégalité $\Theta(u) \leq 0$:

$$\nabla_u L(u^{\text{opt}}, p^{\text{opt}}) = 0 ,$$

$$\Theta(u^{\text{opt}}) \leq 0 \quad , \quad p^{\text{opt}} \geq 0 \quad , \quad \langle p^{\text{opt}}, \Theta(u^{\text{opt}}) \rangle = 0 .$$

Le cas des contraintes égalité est donc très favorable car la résolution des conditions de KKT se résume à la résolution d'un ensemble d'équations.

Conditions de Karush-Kuhn-Tucker

II

Les **gradients partiels** par rapport à u et p du **Lagrangien** L du problème (avec $L(u, p) = J(u) + \langle p, \Theta(u) \rangle$) ont pour expression :

$$\nabla_u L(u, p) = \nabla J(u) + (\Theta'(u))^{\top} \cdot p \quad , \quad \nabla_p L(u, p) = \Theta(u) \text{ .}$$

Dans le cas $U^{\text{ad}} = \mathcal{U}$ (**toutes** les contraintes sont données par Θ), les conditions de KKT prennent les **formes particulières** suivantes :

- pour les **contraintes égalité** $\Theta(u) = 0$:

$$\nabla_u L(u^{\#}, p^{\#}) = 0 \text{ ,}$$

$$\nabla_p L(u^{\#}, p^{\#}) = 0 \text{ ;}$$

- pour les **contraintes inégalité** $\Theta(u) \leq 0$:

$$\nabla_u L(u^{\#}, p^{\#}) = 0 \text{ ,}$$

$$\Theta(u^{\#}) \leq 0 \quad , \quad p^{\#} \geq 0 \quad , \quad \langle p^{\#}, \Theta(u^{\#}) \rangle = 0 \text{ .}$$

Le cas des **contraintes égalité** est donc très **favorable** car la résolution des conditions de KKT se ramène à la résolution d'un ensemble d'équations...

Interprétation marginaliste du multiplicateur

Soit $G : \mathcal{V} \rightarrow \mathbb{R}$ la **fonction de perturbation** du problème (\mathcal{P}_A) , correspondant à la minimisation du problème sous **contraintes perturbées** par une variable $v \in \mathcal{V}$:

$$G(v) = \min_{u \in U^{\text{ad}}} J(u) \quad \text{sous} \quad \Theta(u) - v \in -C .$$

On montre que la fonction G est **sous-différentiable** en $v = 0$, et que le **multiplicateur** $p^\#$ des conditions de **KKT** vérifie :

$$p^\# \in -\partial G(0) .$$

Dans le cas différentiable, $p^\#$ s'interprète donc, au signe près, comme la **sensibilité du coût optimal** par rapport au niveau de contraintes (ici égal à 0). Le lien entre **fonction de perturbation** et **solution du problème** (\mathcal{P}_A) est donné par les deux relations :

$$G(0) = J(u^\#) \quad \text{et} \quad \nabla G(0) = -p^\# .$$

Interprétation marginaliste du multiplicateur

Soit $G : \mathcal{V} \rightarrow \mathbb{R}$ la **fonction de perturbation** du problème (\mathcal{P}_A) , correspondant à la minimisation du problème sous **contraintes perturbées** par une variable $v \in \mathcal{V}$:

$$G(v) = \min_{u \in U^{\text{ad}}} J(u) \quad \text{sous} \quad \Theta(u) - v \in -C .$$

On montre que la fonction G est **sous-différentiable** en $v = 0$, et que le **multiplicateur** p^\sharp des conditions de **KKT** vérifie :

$$p^\sharp \in -\partial G(0) .$$

Dans le cas différentiable, p^\sharp s'interprète donc, au signe près, comme la **sensibilité du coût optimal** par rapport au niveau de contraintes (ici égal à 0). Le lien entre **fonction de perturbation** et **solution du problème** (\mathcal{P}_A) est donné par les deux relations :

$$G(0) = J(u^\sharp) \quad \text{et} \quad \nabla G(0) = -p^\sharp .$$

Lagrangien et point selle

On rappelle que le **Lagrangien** associé au problème (\mathcal{P}_A) est la fonction L , définie sur $U^{\text{ad}} \times C^*$ à valeurs dans \mathbb{R} , définie par :

$$L(u, p) = J(u) + \langle p, \Theta(u) \rangle .$$

On appelle **saut de dualité** la quantité δ (positive ou nulle) :

$$\delta = \min_{u \in U^{\text{ad}}} \max_{p \in C^*} L(u, p) - \max_{p \in C^*} \min_{u \in U^{\text{ad}}} L(u, p) .$$

Le couple $(u^\sharp, p^\sharp) \in U^{\text{ad}} \times C^*$ est un **point selle** de L s'il vérifie :

$$L(u^\sharp, p) \leq L(u^\sharp, p^\sharp) \leq L(u, p^\sharp) , \quad \forall (u, p) \in U^{\text{ad}} \times C^* .$$

On montre que si (u_1^\sharp, p_1^\sharp) et (u_2^\sharp, p_2^\sharp) sont des points selle du Lagrangien L , (u_1^\sharp, p_2^\sharp) et (u_2^\sharp, p_1^\sharp) le sont aussi : l'ensemble S^\sharp des points selle de L est un produit cartésien d'ensembles :

$$S^\sharp = U^\sharp \times P^\sharp .$$

Lagrangien et point selle

On rappelle que le **Lagrangien** associé au problème (\mathcal{P}_A) est la fonction L , définie sur $U^{\text{ad}} \times C^*$ à valeurs dans \mathbb{R} , définie par :

$$L(u, p) = J(u) + \langle p, \Theta(u) \rangle .$$

On appelle **saut de dualité** la quantité δ (positive ou nulle) :

$$\delta = \min_{u \in U^{\text{ad}}} \max_{p \in C^*} L(u, p) - \max_{p \in C^*} \min_{u \in U^{\text{ad}}} L(u, p) .$$

Le couple $(u^\sharp, p^\sharp) \in U^{\text{ad}} \times C^*$ est un **point selle** de L s'il vérifie :

$$L(u^\sharp, p) \leq L(u^\sharp, p^\sharp) \leq L(u, p^\sharp) , \quad \forall (u, p) \in U^{\text{ad}} \times C^* .$$

On montre que si (u_1^\sharp, p_1^\sharp) et (u_2^\sharp, p_2^\sharp) sont des points selle du Lagrangien L , (u_1^\sharp, p_2^\sharp) et (u_2^\sharp, p_1^\sharp) le sont aussi : l'ensemble S^\sharp des points selle de L est un **produit cartésien** d'ensembles :

$$S^\sharp = U^\sharp \times P^\sharp .$$

Lagrangien et point selle



Lorsque le **Lagrangien** L admet un point selle $(u^\#, p^\#)$, le **saut de dualité** est nul et on a les égalités :

$$\max_{p \in C^*} \min_{u \in U^{\text{ad}}} L(u, p) = \min_{u \in U^{\text{ad}}} \max_{p \in C^*} L(u, p) = L(u^\#, p^\#).$$

Les résultats suivants caractérisent l'**optimalité du point selle** :

- (1) Si $(u^\#, p^\#)$ est un **point selle** de L , alors $u^\#$ est **solution** du problème (\mathcal{P}_A) .
- (2) Si J est convexe s.c.i propre coercive, si Θ est C -convexe et continue et si l'hypothèse de **qualification des contraintes** est vérifiée, alors le Lagrangien L admet au moins un **point selle**.

Enfin, si l'on suppose que les fonctions J et Θ sont différentiables, tout **point selle** vérifie les **conditions de KKT**.

Fonction duale et stabilité

On définit la **fonction duale** $H : C^* \rightarrow \mathbb{R}$ par :

$$H(p) = \min_{u \in U^{\text{ad}}} L(u, p),$$

et l'on note $\hat{U}(p)$ l'ensemble (paramétré par p) des valeurs de u réalisant le **minimum** de cette fonction. La **maxi-minimisation** du Lagrangien L est équivalente à la **maximisation** de cette fonction :

$$\max_{p \in C^*} \min_{u \in U^{\text{ad}}} L(u, p) \iff \max_{p \in C^*} H(p),$$

ce qui permet d'obtenir la partie $p^\#$ du **point selle**. La fonction H étant **concave**, un algorithme de type gradient projeté permet en pratique d'effectuer cette maximisation car les contraintes $p \in C^*$ sont suffisamment simples pour que la projection soit facile à faire.

Fonction duale et stabilité



Cependant, pour un $p^\# \in P^\#$, la **minimisation** en u de $L(u, p^\#)$ ne conduit pas forcément à un élément $u^\#$ qui soit la première composante d'un **point selle** du Lagrangien : si l'on a toujours :

$$U^\# \subset \widehat{U}(p^\#),$$

cette **inclusion** peut être **stricte** : l'ensemble $\widehat{U}(p^\#)$ contient alors des **solutions** « **parasites** » n'ayant rien à voir avec les solutions du problème (\mathcal{P}_A).

On introduit la notion de **stabilité** du Lagrangien d'un problème d'optimisation, définie par le fait que l'inclusion ci-dessus est une **égalité** : $U^\# = \widehat{U}(p^\#)$. Alors, le calcul de la première composante du point selle du Lagrangien ne pose plus de difficulté. . .

Considérons le cas J **strictement convexe**. Alors, la minimisation de $L(u, p^\#)$ conduit à une solution $u^\#$ unique. L'ensemble $\widehat{U}(p^\#)$ se réduit donc à un singleton, d'où la **stabilité** du Lagrangien dans ce cas.

Fonction duale et stabilité



Cependant, pour un $p^\# \in P^\#$, la **minimisation** en u de $L(u, p^\#)$ ne conduit pas forcément à un élément $u^\#$ qui soit la première composante d'un **point selle** du Lagrangien : si l'on a toujours :

$$U^\# \subset \hat{U}(p^\#),$$

cette **inclusion** peut être **stricte** : l'ensemble $\hat{U}(p^\#)$ contient alors des **solutions** « **parasites** » n'ayant rien à voir avec les solutions du problème (\mathcal{P}_A).

On introduit la notion de **stabilité** du Lagrangien d'un problème d'optimisation, définie par le fait que l'inclusion ci-dessus est une **égalité** : $U^\# = \hat{U}(p^\#)$. Alors, le calcul de la première composante du point selle du Lagrangien ne pose plus de difficulté. . .

Considérons le cas J **strictement** convexe. Alors, la minimisation de $L(u, p^\#)$ conduit à une solution $u^\#$ **unique**. L'ensemble $\hat{U}(p^\#)$ se réduit donc à un singleton, d'où la **stabilité** du Lagrangien dans ce cas.

Algorithmes de calcul

Plaçons nous dans le cas où la fonction J est **strictement convexe**.
On peut alors montrer que la **fonction duale** H est **différentiable**.
Notant $\hat{u}(p)$ l'unique solution de la minimisation en u de $L(u, p)$,
le gradient de H en p est donné par :

$$\nabla H(p) = \nabla_p L(\hat{u}(p), p) = \Theta(\hat{u}(p)) .$$

La minimisation du **Lagrangien** à p fixé fournit donc la valeur
de la fonction H en p ainsi que le gradient de H en ce point.

La **différentiabilité** de H ouvre la voie aux méthodes de type
gradient pour le calcul de la composante $p^\#$ du point selle.

Algorithmes de calcul



Supposant J **fortement convexe** (de module a) et Θ **Lipschitz** (de constante τ), l'algorithme d'**Uzawa** est un algorithme de **gradient à pas fixe** pour maximiser la fonction H . À l'itération k , partant de $(u^{(k)}, p^{(k)}) \in U^{\text{ad}} \times C^*$, on calcule l'itérée suivant par :

$$u^{(k+1)} = \arg \min_{u \in U^{\text{ad}}} L(u, p^{(k)}) ,$$
$$p^{(k+1)} = \text{proj}_{C^*} (p^{(k)} + \rho \nabla_p L(u^{(k+1)}, p^{(k)})) .$$

ce qui s'écrit de manière équivalente :

$$u^{(k+1)} = \arg \min_{u \in U^{\text{ad}}} (J(u) + \langle p^{(k)}, \Theta(u) \rangle) ,$$
$$p^{(k+1)} = \text{proj}_{C^*} (p^{(k)} + \rho \Theta(u^{(k+1)})) .$$

L'algorithme d'**Uzawa** converge vers l'**unique solution** $u^\#$ du problème (\mathcal{P}_A) avec le choix de pas : $\rho \in]0, \frac{2a}{\tau^2}[$.

Algorithmes de calcul



Une **variante** de l'algorithme d'Uzawa consiste à faire un **pas de gradient** sur la variable u plutôt que la **minimisation complète**.
On obtient alors l'algorithme d'**Arrow–Hurwicz** :

$$\begin{aligned}u^{(k+1)} &= \text{proj}_{U^{\text{ad}}}(u^{(k)} - \epsilon \nabla_u L(u^{(k)}, p^{(k)})) , \\p^{(k+1)} &= \text{proj}_{C^*}(p^{(k)} + \rho \nabla_p L(u^{(k+1)}, p^{(k)})) .\end{aligned}$$

soit encore :

$$\begin{aligned}u^{(k+1)} &= \text{proj}_{U^{\text{ad}}}(u^{(k)} - \epsilon (\nabla J(u^{(k)}) + (\Theta'(u^{(k)}))^{\top} p^{(k)})) , \\p^{(k+1)} &= \text{proj}_{C^*}(p^{(k)} + \rho \Theta(u^{(k+1)})) .\end{aligned}$$

Les conditions de convergence de l'algorithme d'**Arrow–Hurwicz** sont semblables à celles de l'algorithme d'**Uzawa**.

Bibliographie (non exhaustive) de l'optimisation

- **Convexité et optimisation**
G. Cohen, 2019, Cours ENPC.
- **Optimisation différentiable : théorie et algorithmes**
J.-C. Gilbert, 2019, Cours ENSTA.
- **Convex Analysis**
R. T. Rockafellar, 1970, Princeton University Press.
- **Convex Analysis and Variational Problems**
I. Ekeland et R. Temam, 1999, SIAM
- **Convex Analysis and Minimization Algorithms**
J.-B. Hiriart-Urruty et C. Lemaréchal, 1993, Springer.
- **Nonlinear Programming**
D. Bertsekas, 1999, Athena Scientific.
- **Numerical Optimization**
J. Nocedal and S. J. Wright, 1999, Springer.
- **Nonlinear Optimization**
A. Ruszczyński, 2006, Princeton University Press.

- 1 Présentation des grands systèmes
 - Grands systèmes en optimisation
 - Problèmes prototypes
- 2 Rappels d'optimisation convexe
 - Propriétés élémentaires
 - Optimisation sans contrainte explicite
 - Optimisation avec contraintes explicites
- 3 Séance de travaux dirigés
 - Stabilité du Lagrangien
 - Interprétation géométrique de l'algorithme d'Uzawa
 - Optimisation d'un grand réseau de distribution d'eau

- 1 Présentation des grands systèmes
 - Grands systèmes en optimisation
 - Problèmes prototypes
- 2 Rappels d'optimisation convexe
 - Propriétés élémentaires
 - Optimisation sans contrainte explicite
 - Optimisation avec contraintes explicites
- 3 Séance de travaux dirigés
 - Stabilité du Lagrangien
 - Interprétation géométrique de l'algorithme d'Uzawa
 - Optimisation d'un grand réseau de distribution d'eau

Stabilité du Lagrangien

E₁

Soit à résoudre le problème de minimisation suivant :

$$\min_{u \in [-1,1]} -u \quad \text{sous } u \leq 0 ,$$

dont la solution (évidente) est $u^\# = 0$.

- 1 Montrer que le **Lagrangien** associé à ce problème admet sur $[-1, 1] \times \mathbb{R}_+$ un unique point selle $(u^\#, p^\#)$.
- 2 Calculer le minimum et l'argmin en u du Lagrangien au point $p = p^\#$ (algorithme d'**Uzawa**). Qu'en déduit-on ?

Stabilité du Lagrangien

R₁

On vérifie aisément que l'**unique** solution du problème est $u^\# = 0$.
Le **Lagrangien** du problème, défini sur $[-1, 1] \times \mathbb{R}_+$, est

$$L(u, p) = -u + pu = (p - 1)u .$$

La dérivée en u de L (égale à $p - 1$) ne s'annule qu'en $p^\# = 1$. Par **KKT**, on déduit que l'unique **point selle** de L est $(u^\#, p^\#) = (0, 1)$.

En $p^\# = 1$, le Lagrangien L est **nul**, et on a donc $\widehat{U}(p^\#) = [-1, 1]$,
alors que l'ensemble des solutions du problème est $U^\# = \{0\}$:

$$U^\# \subsetneq \widehat{U}(p^\#) .$$

Le **Lagrangien** de ce problème n'est donc pas **stable**.

Stabilité du Lagrangien

R₂

Le calcul de la **fonction duale** H conduit à

$$H(p) = \min_{u \in [-1,1]} (p-1)u = \begin{cases} p-1 & \text{si } p \in [0, 1], \\ 1-p & \text{si } p \geq 1, \end{cases}$$

Durant les itérations de l'algorithme d'**Uzawa**,

- si $p^{(k)} < 1$, la minimisation en u donne $u^{(k+1)} = 1$ et la mise à jour du multiplicateur fait **augmenter** $p^{(k)}$,
- sinon, $p^{(k)} > 1$, la minimisation en u donne $u^{(k+1)} = -1$ et la mise à jour du multiplicateur fait **diminuer** $p^{(k)}$.

La suite des solutions $\{u^{(k)}\}_{k \in \mathbb{N}}$ va donc osciller entre les valeurs $+1$ et -1 , alors que la solution optimale est $u^\# = 0 \dots$

- 1 Présentation des grands systèmes
 - Grands systèmes en optimisation
 - Problèmes prototypes
- 2 Rappels d'optimisation convexe
 - Propriétés élémentaires
 - Optimisation sans contrainte explicite
 - Optimisation avec contraintes explicites
- 3 Séance de travaux dirigés
 - Stabilité du Lagrangien
 - **Interprétation géométrique de l'algorithme d'Uzawa**
 - Optimisation d'un grand réseau de distribution d'eau

Interprétation géométrique de l'algorithme d'Uzawa

E₁

Soit le problème d'optimisation sous contraintes d'égalité :

$$\min_{u \in U^{\text{ad}}} J(u) \quad \text{sous} \quad \Theta(u) = 0 ,$$

que l'on veut résoudre par l'algorithme d'Uzawa.

- 1 Réécrire l'algorithme d'Uzawa à l'aide de la fonction de perturbation G en introduisant une variable supplémentaire v .
- 2 Interpréter de façon géométrique ce nouvel algorithme.
- 3 Que se passe t'il dans le cas où la fonction G n'est pas strictement convexe.
- 4 Peut-on généraliser au cas des contraintes de type inégalité ?

Interprétation géométrique de l'algorithme d'Uzawa

R₁

L'itération k de l'algorithme d'Uzawa s'écrit :

$$\begin{aligned} u^{(k+1)} &\in \arg \min_{u \in U^{\text{ad}}} J(u) + \langle \lambda^{(k)}, \Theta(u) \rangle, \\ \lambda^{(k+1)} &= \lambda^{(k)} + \rho \Theta(u^{(k+1)}). \end{aligned}$$

L'étape de minimisation se réécrit de manière équivalente :

$$\min_{v \in \mathcal{V}} \min_{u \in U^{\text{ad}}} J(u) + \langle \lambda^{(k)}, v \rangle \quad \text{s.t.} \quad \Theta(u) - v = 0.$$

Utilisant la **fonction de perturbation** G :

$$G(v) = \min_{u \in U^{\text{ad}}} J(u) \quad \text{sous} \quad \Theta(u) - v = 0,$$

cette étape de minimisation devient :

$$\min_{v \in \mathcal{V}} G(v) + \langle \lambda^{(k)}, v \rangle.$$

Interprétation géométrique de l'algorithme d'Uzawa R₂

Cette réécriture permet donc de mettre l'algorithme d'Uzawa sous la forme :

$$\begin{aligned}v^{(k+1)} &\in \arg \min_{v \in \mathcal{V}} G(v) + \langle \lambda^{(k)}, v \rangle, \\ \lambda^{(k+1)} &= \lambda^{(k)} + \rho v^{(k+1)}.\end{aligned}$$

On rappelle que le problème consiste à calculer $G(0)$.

Du point de vue géométrique, l'algorithme consiste donc :

- Étape 1 : minimiser l'écart entre $G(\cdot)$ et $\langle -\lambda^{(k)}, \cdot \rangle$.
- Étape 2 : modifier la pente $-\lambda^{(k)}$ si $v^{(k+1)} \neq 0$.

Cette interprétation est conceptuelle car G n'est pas connue...

Interprétation géométrique de l'algorithme d'Uzawa R₂

Cette réécriture permet donc de mettre l'algorithme d'Uzawa sous la forme :

$$\begin{aligned}v^{(k+1)} &\in \arg \min_{v \in \mathcal{V}} G(v) + \langle \lambda^{(k)}, v \rangle, \\ \lambda^{(k+1)} &= \lambda^{(k)} + \rho v^{(k+1)}.\end{aligned}$$

On rappelle que le problème consiste à calculer $G(0)$.

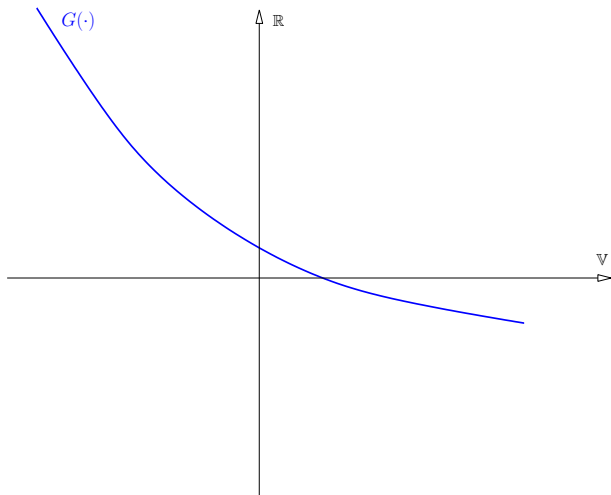
Du point de vue **géométrique**, l'algorithme consiste donc :

- **Étape 1** : **minimiser** l'écart entre $G(\cdot)$ et $\langle -\lambda^{(k)}, \cdot \rangle$.
- **Étape 2** : **modifier** la pente $-\lambda^{(k)}$ si $v^{(k+1)} \neq 0$.

Cette interprétation est **conceptuelle** car G n'est pas connue...

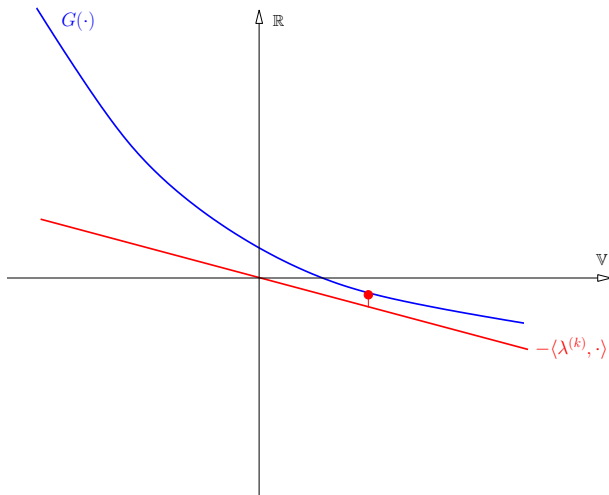
Interprétation géométrique de l'algorithme d'Uzawa

R_3



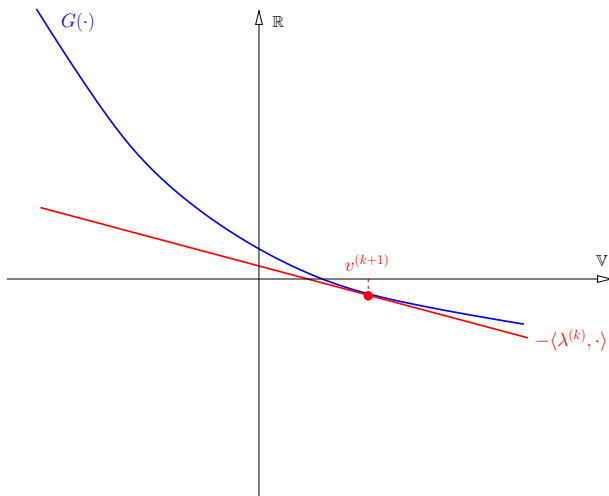
Interprétation géométrique de l'algorithme d'Uzawa

R₃



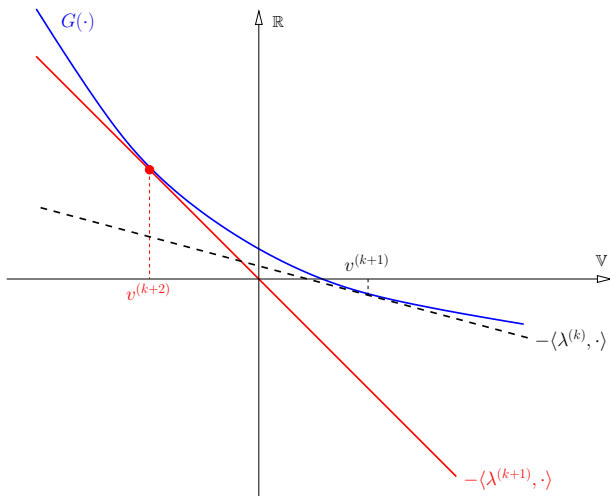
Interprétation géométrique de l'algorithme d'Uzawa

R₃



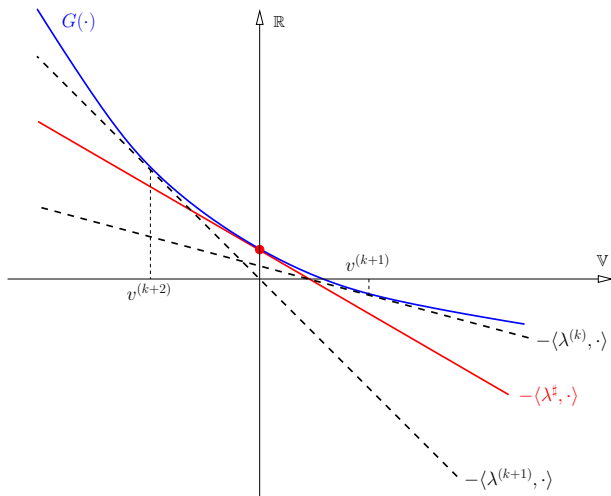
Interprétation géométrique de l'algorithme d'Uzawa

R₃



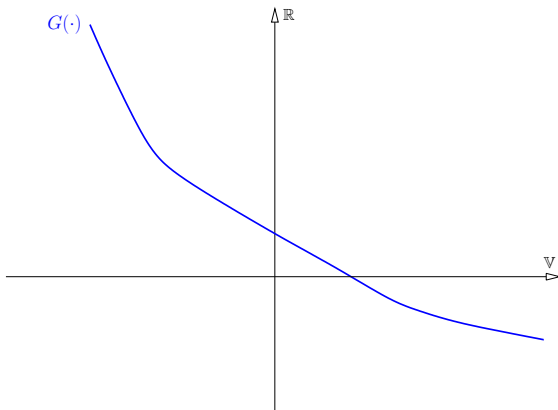
Interprétation géométrique de l'algorithme d'Uzawa

R₃



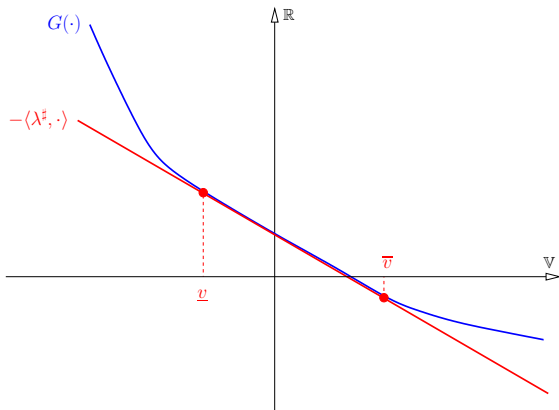
Interprétation géométrique de l'algorithme d'Uzawa

R₄



Interprétation géométrique de l'algorithme d'Uzawa

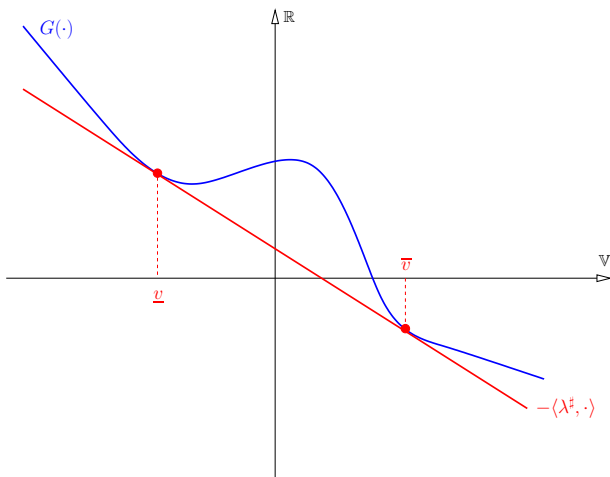
R₄



Même si $\{\lambda^{(k)}\}_{k \in \mathbb{N}}$ converge vers $\lambda^\#$, le niveau de contrainte $v^{(k)}$ oscille entre \underline{v} et \bar{v} , et la valeur $v^\# = 0$ n'est **jamais obtenue**.

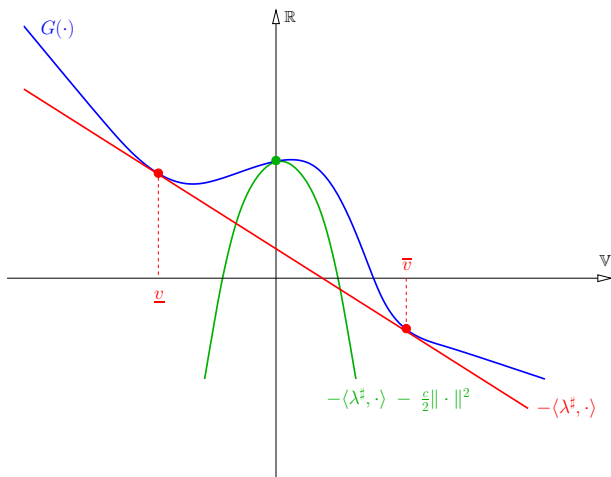
Interprétation géométrique de l'algorithme d'Uzawa

R₅



Interprétation géométrique de l'algorithme d'Uzawa

R₅

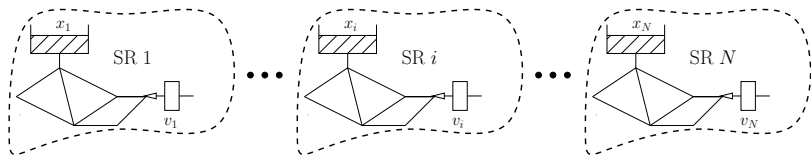


Dans le cas **non convexe**, utiliser le **Lagrangien augmenté**!

- 1 Présentation des grands systèmes
 - Grands systèmes en optimisation
 - Problèmes prototypes
- 2 Rappels d'optimisation convexe
 - Propriétés élémentaires
 - Optimisation sans contrainte explicite
 - Optimisation avec contraintes explicites
- 3 Séance de travaux dirigés
 - Stabilité du Lagrangien
 - Interprétation géométrique de l'algorithme d'Uzawa
 - Optimisation d'un grand réseau de distribution d'eau

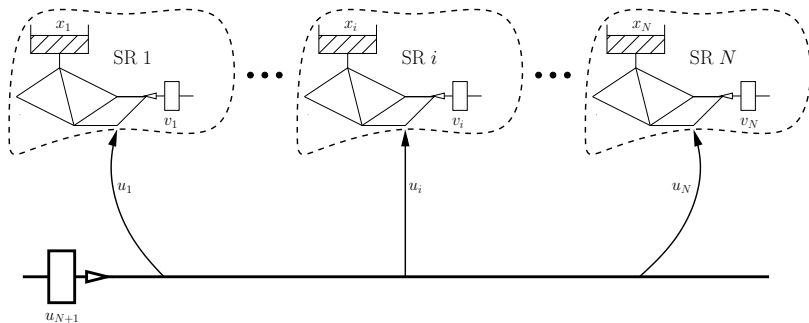
Réseau de distribution d'eau

E_0



Réseau de distribution d'eau

E_0



Réseau de distribution d'eau

E₁

On s'intéresse à l'optimisation du coût de fonctionnement d'un **réseau de distribution d'eau** sur une période de **24** heures. Cette période est scindée en **2** pas de temps (jour/nuit) correspondant aux variations du coût de l'énergie.

Ce réseau est constitué de N sous-réseaux **indépendants**. Chaque sous-réseau i comporte une **usine** injectant la quantité d'eau $v_{i,t}$ au pas de temps t pour un coût $\frac{1}{2}a_{i,t}v_{i,t}^2$, et un **réservoir** : on impose que le niveau du réservoir $x_{i,t}$ reste positif à $t = 1$, et revienne à son niveau initial $x_{i,0}$ à $t = 2$. Notant $c_{i,t}$ la **consommation** en eau au pas de temps t , la dynamique du réservoir s'écrit :

$$x_{i,1} = x_{i,0} + v_{i,1} - c_{i,1} \geq 0 \quad , \quad x_{i,2} = x_{i,1} + v_{i,2} - c_{i,2} = x_{i,0} \quad ,$$

d'où des **contraintes** sur les injections d'eau de la forme :

$$v_{i,1} \geq \bar{v}_{i,1} \quad , \quad v_{i,1} + v_{i,2} = \bar{v}_{i,2} \quad .$$

Pour des raisons de **cohérence**, on suppose que $\bar{v}_{i,2} \geq \bar{v}_{i,1} \geq 0$.

Réseau de distribution d'eau

E₂

Le **problème d'optimisation** associé au sous-réseau i s'écrit alors :

$$\min_{(v_{i,1}, v_{i,2})} \frac{1}{2} (a_{i,1} v_{i,1}^2 + a_{i,2} v_{i,2}^2) ,$$

sous les contraintes

$$v_{i,1} \geq \bar{v}_{i,1} ,$$

$$v_{i,1} + v_{i,2} - \bar{v}_{i,2} = 0 .$$

Pour **optimiser le réseau global**, il suffit de résoudre le problème de minimisation associé à chacun des sous-réseaux, tous **indépendants**.

Réseau de distribution d'eau

E₃

Pour des raisons de sécurité, on construit une nouvelle usine alimentant **simultanément** les N sous-réseaux. Notant $u_{i,t}$ la quantité d'eau fournie à t par cette usine au sous-réseau i , le problème d'optimisation du sous-réseau i devient :

$$\min_{(v_{i,1}, v_{i,2})} \frac{1}{2} (a_{i,1} v_{i,1}^2 + a_{i,2} v_{i,2}^2) ,$$

sous les contraintes

$$u_{i,1} + v_{i,1} \geq \bar{v}_{i,1} ,$$

$$u_{i,1} + u_{i,2} + v_{i,1} + v_{i,2} - \bar{v}_{i,2} = 0 .$$

Le **résultat** de la minimisation en $(v_{i,1}, v_{i,2})$ de ce problème est une fonction des variables $(u_{i,1}, u_{i,2})$, que l'on note $J_i(u_{i,1}, u_{i,2})$.

Réseau de distribution d'eau

E₄

La **quantité globale d'eau** fournie au pas de temps t par la nouvelle usine est égale à $\sum_{i=1}^N u_{i,t}$. Notant J_{N+1} la fonction de coût de cette usine, le **problème global** de l'optimisation du réseau est :

$$\min_{\{(u_{i,1}, u_{i,2})\}_{i=1}^N} \sum_{i=1}^N J_i(u_{i,1}, u_{i,2}) + J_{N+1} \left(\sum_{i=1}^N u_{i,1}, \sum_{i=1}^N u_{i,2} \right),$$

soit, avec les **variables supplémentaires** $u_{N+1,t} = \sum_{i=1}^N u_{i,t}$,

$$\min_{\{(u_{i,1}, u_{i,2})\}_{i=1}^{N+1}} \sum_{i=1}^{N+1} J_i(u_{i,1}, u_{i,2}),$$

sous les contraintes

$$\sum_{i=1}^N u_{i,t} - u_{N+1,t} = 0, \quad t = 1, 2.$$

Réseau de distribution d'eau

E₅

On s'intéresse au **problème d'optimisation** du sous-réseau i

$$\min_{(v_{i,1}, v_{i,2})} \frac{1}{2} (a_{i,1} v_{i,1}^2 + a_{i,2} v_{i,2}^2) ,$$

sous les contraintes

$$v_{i,1} \geq \bar{v}_{i,1} ,$$

$$v_{i,1} + v_{i,2} - \bar{v}_{i,2} = 0 .$$

- 1 Mettre ce problème sous **forme standard**, s'assurer de l'**existence** et de l'**unicité** de la solution et **calculer** cette solution en résolvant directement les **conditions de KKT**.
- 2 Formuler le problème de minimisation que l'on obtient, **avant toute opération de minimisation**, lorsque l'on **connecte** les N réseaux par la nouvelle usine commune ; expliquer pourquoi la méthode de résolution par KKT devient **impraticable**.

Réseau de distribution d'eau

R₁

Le problème d'optimisation s'écrit sous la **forme standard** :

$$\min_{(v_{i,1}, v_{i,2}) \in \mathbb{R}^2} \frac{1}{2} (a_{i,1} v_{i,1}^2 + a_{i,2} v_{i,2}^2) ,$$

sous les contraintes

$$\bar{v}_{i,1} - v_{i,1} \leq 0 ,$$

$$v_{i,1} + v_{i,2} - \bar{v}_{i,2} = 0 .$$

et vérifie toutes les **propriétés** requises :

- l'ensemble admissible $U^{\text{ad}} = \mathbb{R}^2$ est **convexe fermé**,
- les coefficients $a_{i,1}$ et $a_{i,2}$ étant strictement positifs, la fonction coût J est fortement convexe (donc coercive), continue, à gradient linéaire (donc Lipschitzien),
- le cône des contraintes C est égal à $\mathbb{R}_+ \times \{0\}$,
- les contraintes Θ sont linéaires (donc convexes sur le cône C et continues), et les hypothèses de qualification sont vérifiées.

Réseau de distribution d'eau

R₂

Le **Lagrangien** du problème a pour expression :

$$\frac{1}{2} (a_{i,1} v_{i,1}^2 + a_{i,2} v_{i,2}^2) + \lambda_{i,1} (\bar{v}_{i,1} - v_{i,1}) + \lambda_{i,2} (v_{i,1} + v_{i,2} - \bar{v}_{i,2}) ,$$

et les conditions de KKT s'écrivent :

$$a_{i,1} v_{i,1} - \lambda_{i,1} + \lambda_{i,2} = 0 ,$$

$$a_{i,2} v_{i,2} + \lambda_{i,2} = 0 ,$$

$$\bar{v}_{i,1} - v_{i,1} \leq 0 \quad , \quad \lambda_{i,1} \geq 0 ,$$

$$\lambda_{i,1} (\bar{v}_{i,1} - v_{i,1}) = 0 ,$$

$$v_{i,1} + v_{i,2} - \bar{v}_{i,2} = 0 \quad , \quad \lambda_{i,2} \text{ quelconque .}$$

Pour résoudre ce **système d'inéquations**, on **joue aux devinettes** ...

Réseau de distribution d'eau

R₃

- Si la **contrainte inégalité n'est pas saturée** : $\bar{v}_{i,1} - v_{i,1} < 0$, la condition des **écarts complémentaires** implique $\lambda_{i,1} = 0$. Les conditions de **KKT** se réduisent alors à un **système linéaire**, dont la solution

$$v_{i,1} = \frac{a_{i,2}}{a_{i,1} + a_{i,2}} \bar{v}_{i,2}, \quad v_{i,2} = \frac{a_{i,1}}{a_{i,1} + a_{i,2}} \bar{v}_{i,2}, \quad \lambda_{i,1} = 0, \quad \lambda_{i,2} = -\frac{a_{i,1}a_{i,2}}{a_{i,1} + a_{i,2}} \bar{v}_{i,2}$$

est solution du problème **pourvu que** la contrainte inégalité ne soit pas saturée, et donc si $(a_{i,1} + a_{i,2})\bar{v}_{i,1} - a_{i,2}\bar{v}_{i,2} < 0$.

- Si la **contrainte inégalité est saturée** : $v_{i,1} = \bar{v}_{i,1}$, les conditions de **KKT** fournissent un autre **système linéaire** dont la solution

$$v_{i,1} = \bar{v}_{i,1}, \quad v_{i,2} = \bar{v}_{i,2} - \bar{v}_{i,1}, \quad \lambda_{i,1} = (a_{i,1} + a_{i,2})\bar{v}_{i,1} - a_{i,2}\bar{v}_{i,2}, \quad \lambda_{i,2} = -a_{i,2}(\bar{v}_{i,2} - \bar{v}_{i,1})$$

est solution du problème **pourvu que** l'on ait $\lambda_{i,1} \geq 0$, et donc sous la condition : $(a_{i,1} + a_{i,2})\bar{v}_{i,1} - a_{i,2}\bar{v}_{i,2} \geq 0$.

On a donc obtenu la **solution du problème** dans tous les cas.

Réseau de distribution d'eau

R₄

Connectant les N sous-réseaux par l'usine commune, le problème d'optimisation **global** à résoudre s'écrit :

$$\min_{\{(u_{i,1}, u_{i,2})\}_{i=1}^N} \sum_{i=1}^N J_i(u_{i,1}, u_{i,2}) + J_{N+1} \left(\sum_{i=1}^N u_{i,1}, \sum_{i=1}^N u_{i,2} \right),$$

avec, pour $i = 1, \dots, N$,

$$J_i(u_{i,1}, u_{i,2}) = \min_{(v_{i,1}, v_{i,2})} \frac{1}{2} (a_{i,1} v_{i,1}^2 + a_{i,2} v_{i,2}^2),$$

$$\text{sous } u_{i,1} + v_{i,1} \geq \bar{v}_{i,1},$$

$$u_{i,1} + u_{i,2} + v_{i,1} + v_{i,2} - \bar{v}_{i,2} = 0.$$

et

$$J_{N+1}(u_{N+1,1}, u_{N+1,2}) = \frac{1}{2} (a_{N+1,1} u_{N+1,1}^2 + a_{N+1,2} u_{N+1,2}^2).$$

Réseau de distribution d'eau

R₅

Avant toute opération de minimisation, le problème **global** s'écrit :

$$\min_{\{(u_{i,1}, u_{i,2}, v_{i,1}, v_{i,2})\}_{i=1}^N} \frac{1}{2} \sum_{i=1}^N (a_{i,1} v_{i,1}^2 + a_{i,2} v_{i,2}^2) + \frac{1}{2} \left(a_{N+1,1} \left(\sum_{i=1}^N u_{i,1} \right)^2 + a_{N+1,2} \left(\sum_{i=1}^N u_{i,2} \right)^2 \right),$$

$$\text{sous } \bar{v}_{i,1} - u_{i,1} - v_{i,1} \leq 0, \quad i = 1, \dots, N,$$

$$u_{i,1} + u_{i,2} + v_{i,1} + v_{i,2} - \bar{v}_{i,2} = 0, \quad i = 1, \dots, N.$$

Pour calculer par **KKT** la solution de ce problème, on doit prendre en compte **N contraintes inégalité**, donc considérer **2^N alternatives** dans le jeu des devinettes issu des conditions de KKT.

Pour un grand réseau en Île de France, **N** est de l'ordre de **40** et le nombre d'alternatives de l'ordre de **10^{12}** . La résolution consiste à former et résoudre plus de **10^{12}** systèmes linéaires de dimension **$6N$** , ce qui est lourd...