# Méthode des centres et de barrière: aspects asymptotiques

Jean-Pierre Dussault

Départment d'informatique
Université de Sherbrooke
email: Jean-Pierre.Dussault@USherbrooke.CA

Journée JFRO 25 novembre 2008

# Plan

1. Introduction
   - Centre
   - Équations
   - Barrière
   - Karmarkar
2. Trajectoires différentiables
   - Analyse asymptotique des trajectoires
   - Implications algorithmiques
   - Poursuite de la trajectoire
3. Conclusion et perspectives

# Plan

# Formulation

$$\begin{aligned} f^* = f(x^*) = \min \quad & f(x) \\ \text{sujet à} \quad & g(x) \leq 0 \end{aligned}$$

Les fonctions $f$ et $g$ sont suffisamment différentiables.

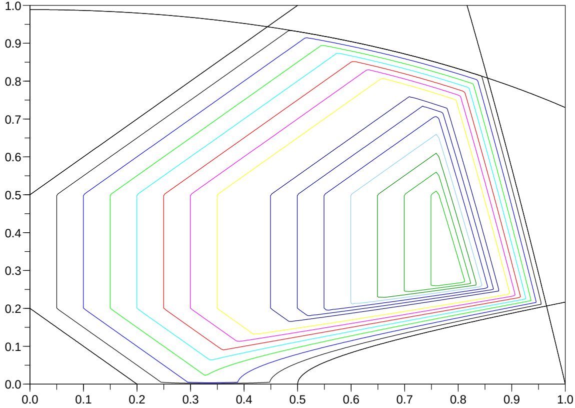
$$E = \{x : g(x) \leq 0\}$$

# Centre de $E$

La fonction $c(x) = \max g_i(x)$.

# Centre de $E$

La fonction $c(x) = \max g_i(x)$.

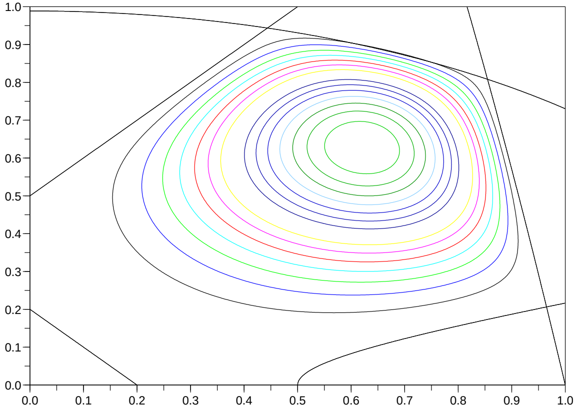


# Centre de $E$

Version lisse : la fonction $c(x) = \prod_i (-g_i(x)) = e^{\sum \log(-g_i(x))}$.

# Centre de $E$

Version lisse : la fonction $c(x) = \prod_i (-g_i(x)) = e^{\sum \log(-g_i(x))}$.

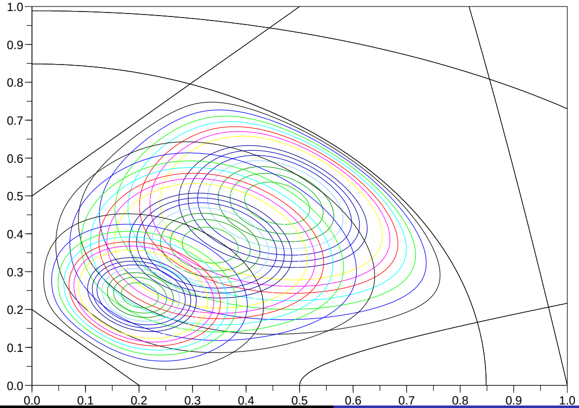


# Centre pondéré de $E \cap \{x : f(x) \leq f(x_0)\}$

Ajoutons la contrainte $(\overline{f_0} - f(x))^p \geq 0$ pour $\overline{f_0} = f(x_0), x_0 = (0.6, 0.6),\ p = 1, 4, 16.$

# Centre pondéré de $E \cap \{x : f(x) \leq f(x_0)\}$

Ajoutons la contrainte $(\overline{f_0} - f(x))^p \geq 0$ pour $\overline{f_0} = f(x_0), x_0 = (0.6, 0.6),\ p = 1, 4, 16.$

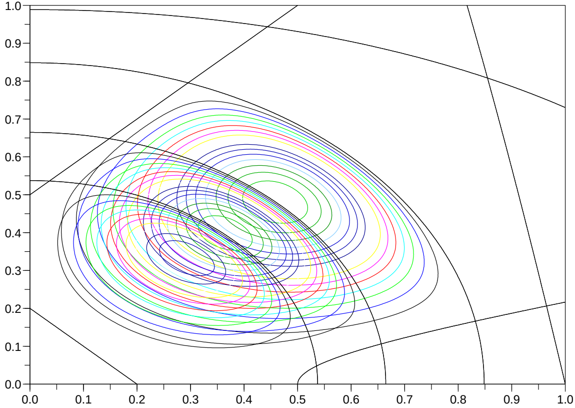


# Centre de $E \cap \{x : f(x) \leq \overline{f_0}\}$ pour choix de $x_0$

Ajoutons la contrainte $(\overline{f_0} - f(x)) \geq 0$ pour $x_0 = [0.6, 0.47, 0.37]e$.

# Centre de $E \cap \{x : f(x) \leq \overline{f_0}\}$ pour choix de $x_0$

Ajoutons la contrainte $(\overline{f_0} - f(x)) \geq 0$ pour $x_0 = [0.6, 0.47, 0.37]e$.



# Équations primales

Prenons le log de la fonction lisse à maximiser.

$$\max_{g_i(x)<0, f(x)<\overline{f_0}} \overline{\phi}(x, p, \overline{f_0}) = \left\{ p \log(\overline{f_0} - f(x)) + \sum \log(-g_i(x)) \right\}$$

Un point stationnaire satisfait à :

$$-\nabla f(x) + \sum \frac{\overline{f_0} - f(x)}{p g_i(x)} \nabla g_i(x) = 0.$$

Comme Mifflin, $\rho = 1/p$ :

$$-\nabla f(x) + \sum \rho \frac{\overline{f_0} - f(x)}{g_i(x)} \nabla g_i(x) = 0.$$

# Équations primales

Prenons le log de la fonction lisse à maximiser.

$$\max_{g_i(x)<0, f(x)<\overline{f_0}} \overline{\phi}(x,p,\overline{f_0}) = \left\{ p \log(\overline{f_0} - f(x)) + \sum \log(-g_i(x)) \right\}$$

Un point stationnaire satisfait à :

$$-\nabla f(x) + \sum \frac{\overline{f_0} - f(x)}{p g_i(x)} \nabla g_i(x) = 0.$$

Comme Mifflin, $\rho = 1/p$ :

$$-\nabla f(x) + \sum \rho \frac{\overline{f_0} - f(x)}{g_i(x)} \nabla g_i(x) = 0.$$

# Équations primales

Prenons le log de la fonction lisse à maximiser.

$$\max_{g_i(x)<0, f(x)<\overline{f_0}} \overline{\phi}(x, p, \overline{f_0}) = \left\{ p \log(\overline{f_0} - f(x)) + \sum \log(-g_i(x)) \right\}$$

Un point stationnaire satisfait à :

$$-\nabla f(x) + \sum \frac{\overline{f_0} - f(x)}{p g_i(x)} \nabla g_i(x) = 0.$$

Comme Mifflin, $\rho = 1/p$ :

$$-\nabla f(x) + \sum \rho \frac{\overline{f_0} - f(x)}{g_i(x)} \nabla g_i(x) = 0.$$

UNIVERSITÉ DE SHERBROOKE

# Équations primales-duales

Utilisons la substitution

$$\lambda_i = \rho \frac{\overline{f_0} - f(x)}{g_i(x)}.$$

Récrivons la condition de stationnarité :

$$\begin{aligned} -\nabla f(x) + \sum \lambda_i \nabla g_i(x) &= 0 \\ \lambda_i g_i(x) - \rho(\overline{f_0} - f(x)) &= 0. \end{aligned}$$

On voit que $\lambda_i g_i(x)$ peut être amené à zéro de deux manières, soit $\rho \searrow 0$, soit $\overline{f_0} \searrow f(x^*)$.

# Équations primales-duales

Utilisons la substitution

$$\lambda_i = \rho \frac{\overline{f_0} - f(x)}{g_i(x)}.$$

Récrivons la condition de stationnarité :

$$\begin{aligned} -\nabla f(x) + \sum \lambda_i \nabla g_i(x) &= 0 \\ \lambda_i g_i(x) - \rho(\overline{f_0} - f(x)) &= 0. \end{aligned}$$

On voit que $\lambda_i g_i(x)$ peut être amené à zéro de deux manières, soit $\rho \searrow 0$, soit $\overline{f_0} \searrow f(x^*)$.

UNIVERSITÉ DE SHERBROOKE

# Équations primales-duales

Utilisons la substitution

$$\lambda_i = \rho \frac{\overline{f_0} - f(x)}{g_i(x)}.$$

Récrivons la condition de stationnarité :

$$\begin{array}{rcl} -\nabla f(x) + \sum \lambda_i \nabla g_i(x) & = & 0 \\ \lambda_i g_i(x) - \rho(\overline{f_0} - f(x)) & = & 0. \end{array}$$

On voit que $\lambda_i g_i(x)$ peut être amené à zéro de deux manières, soit $\rho \searrow 0$, soit $\overline{f_0} \searrow f(x^*)$.

# Barrières

Si on choisit $\rho \searrow 0$, on peut bien se priver de $\overline{f_0} - f(x)$, et obtenir les équations de la barrière logarithmique :

$$\begin{aligned} -\nabla f(x) + \sum \lambda_i \nabla g_i(x) &= 0 \\ \lambda_i g_i(x) - \rho &= 0. \end{aligned}$$

Défaisons la substitution $\lambda_i = \frac{\rho}{g_i(x)}$, et on obtient

$$-\nabla f(x) + \sum \frac{\rho}{g_i(x)} \nabla g_i(x) = 0$$

qui n'est rien d'autre que $\nabla_x \phi(x, \rho)$ :

$$\phi(x, \rho) = -f(x) + \rho \sum \log(-g_i(x)).$$

# Barrières

Si on choisit $\rho \searrow 0$, on peut bien se priver de $\overline{f_0} - f(x)$, et obtenir les équations de la barrière logarithmique :

$$\begin{aligned} -\nabla f(x) + \sum \lambda_i \nabla g_i(x) &= 0 \\ \lambda_i g_i(x) - \rho &= 0. \end{aligned}$$

Défaisons la substitution $\lambda_i = \frac{\rho}{g_i(x)}$, et on obtient

$$-\nabla f(x) + \sum \frac{\rho}{g_i(x)} \nabla g_i(x) = 0$$

qui n'est rien d'autre que $\nabla_x \phi(x, \rho)$ :

$$\phi(x, \rho) = -f(x) + \rho \sum \log(-g_i(x)).$$

# Barrières

Si on choisit $\rho \searrow 0$, on peut bien se priver de $\overline{f_0} - f(x)$, et obtenir les équations de la barrière logarithmique :

$$\begin{aligned} -\nabla f(x) + \sum \lambda_i \nabla g_i(x) &= 0 \\ \lambda_i g_i(x) - \rho &= 0. \end{aligned}$$

Défaisons la substitution $\lambda_i = \frac{\rho}{g_i(x)}$, et on obtient

$$-\nabla f(x) + \sum \frac{\rho}{g_i(x)} \nabla g_i(x) = 0$$

qui n'est rien d'autre que $\nabla_x \phi(x, \rho)$ :

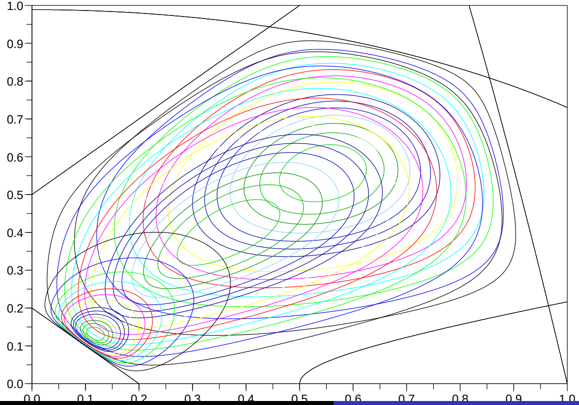
$$\phi(x, \rho) = -f(x) + \rho \sum \log(-g_i(x)).$$

# Une variante — Karmarkar

Ajoutons la contrainte $(f(x) - \underline{f_0})^p \geq 0$ pour $\underline{f_0} < f^*$. Sur le graphe, $\underline{f_0} = 0$.

# Une variante — Karmarkar

Ajoutons la contrainte $(f(x) - \underline{f_0})^p \geq 0$ pour $\underline{f_0} < f^*$. Sur le graphe, $\underline{f_0} = 0$.



# Équations primales-duales, résumé

$$-\nabla f(x) + \sum \lambda_i \nabla g_i(x) = 0$$

$\boxed{\text{Centre :}}$

$$\lambda_i g_i(x) - \rho(\overline{f_0} - f(x)) = 0 \qquad \text{avec } \overline{f_0} > f^*$$

$\boxed{\text{Barrière :}}$

$$\lambda_i g_i(x) - \rho = 0$$

$\boxed{\text{Karmarkar :}}$

$$\lambda_i g_i(x) - \rho(f(x) - \underline{f_0}) = 0 \qquad \text{avec } \underline{f_0} < f^*$$

On voit que $\lambda_i g_i(x)$ peut être amené à zéro de trois manières, soit $\rho \searrow 0$, soit $\overline{f_0} \searrow f(x^*)$, ou $\underline{f_0} \nearrow f(x^*)$.

# Équations primales-duales, résumé

$$-\nabla f(x) + \sum \lambda_i \nabla g_i(x) = 0$$

$\boxed{\text{Centre :}}$

$$\lambda_i g_i(x) - \rho(\overline{f_0} - f(x)) = 0 \qquad \text{avec } \overline{f_0} > f^*$$

$\boxed{\text{Barrière :}}$

$$\lambda_i g_i(x) - \rho = 0$$

$\boxed{\text{Karmarkar :}}$

$$\lambda_i g_i(x) - \rho(f(x) - \underline{f_0}) = 0 \qquad \text{avec } \underline{f_0} < f^*$$

On voit que $\lambda_i g_i(x)$ peut être amené à zéro de trois manières, soit $\rho \searrow 0$, soit $\overline{f_0} \searrow f(x^*)$, ou $\underline{f_0} \nearrow f(x^*)$.

# Équations primales-duales, résumé

$$-\nabla f(x) + \sum \lambda_i \nabla g_i(x) = 0$$

Centre :

$$\lambda_i g_i(x) - \rho(\overline{f_0} - f(x)) = 0 \qquad \text{avec } \overline{f_0} > f^*$$

Barrière :

$$\lambda_i g_i(x) - \rho = 0$$

Karmarkar :

$$\lambda_i g_i(x) - \rho(f(x) - \underline{f_0}) = 0 \qquad \text{avec } \underline{f_0} < f^*$$

On voit que $\lambda_i g_i(x)$ peut être amené à zéro de trois manières, soit $\rho \searrow 0$, soit $\overline{f_0} \searrow f(x^*)$, ou $\underline{f_0} \nearrow f(x^*)$.

# Équations primales-duales, résumé

$$
\begin{array}{rcll}
-\nabla f(x) + \sum \lambda_i \nabla g_i(x) & = & 0 & \\
\boxed{\text{Centre :}} & & & \\
\lambda_i g_i(x) - \rho(\overline{f_0} - f(x)) & = & 0 & \text{avec } \overline{f_0} > f^* \\
\boxed{\text{Barrière :}} & & & \\
\lambda_i g_i(x) - \rho & = & 0 & \\
\boxed{\text{Karmarkar :}} & & & \\
\lambda_i g_i(x) - \rho(f(x) - \underline{f_0}) & = & 0 & \text{avec } \underline{f_0} < f^*
\end{array}
$$

On voit que $\lambda_i g_i(x)$ peut être amené à zéro de trois manières, soit $\rho \searrow 0$, soit $\overline{f_0} \searrow f(x^*)$, ou $\underline{f_0} \nearrow f(x^*)$.

# Équations primales-duales, résumé

$$
\begin{array}{rcll}
-\nabla f(x) + \sum \lambda_i \nabla g_i(x) & = & 0 & \\
\boxed{\text{Centre :}} & & & \\
\lambda_i g_i(x) - \rho(\overline{f_0} - f(x)) & = & 0 & \text{avec } \overline{f_0} > f^* \\
\boxed{\text{Barrière :}} & & & \\
\lambda_i g_i(x) - \rho & = & 0 & \\
\boxed{\text{Karmarkar :}} & & & \\
\lambda_i g_i(x) - \rho(f(x) - \underline{f_0}) & = & 0 & \text{avec } \underline{f_0} < f^*
\end{array}
$$

On voit que $\lambda_i g_i(x)$ peut être amené à zéro de trois manières, soit $\rho \searrow 0$, soit $\overline{f_0} \searrow f(x^*)$, ou $\underline{f_0} \nearrow f(x^*)$.

# Plan

# Paramétrisations de la trajectoire centrale

## Hypothèses

- Indépendance linéaire des contraintes actives (LICQ) ;
- complémentarité stricte (CS) ; $|\{i : \lambda_i > 0\}| = q$ ;
- conditions suffisantes de second ordre en $x^*$ (SOSC).

## Théorème

- *[trajectoire]* $x(\rho)$ *différentiable en* $\rho = 0$ ;
  $x(\underline{f_0})$ *et* $x(\overline{f_0})$ *différentiables en* $f_0 = f^*$ ;
- *[barrière]* $x(\rho)$ *minimiseur de* $\phi(x, \rho)$ ;
- *[centre]* $x(\overline{f_0})$ *minimiseur de* $\bar{\phi}(x, p, \overline{f_0})$ ;
- *[Karmarkar]* $p > q$ *assure que* $x(\underline{f_0})$ *minimiseur de* $\underline{\phi}(x, p, \underline{f_0})$.

# Paramétrisations de la trajectoire centrale

## Hypothèses

- Indépendance linéaire des contraintes actives (LICQ) ;
- complémentarité stricte (CS) ; $|\{i : \lambda_i > 0\}| = q$ ;
- conditions suffisantes de second ordre en $x^*$ (SOSC).

## Théorème

- *[trajectoire]* $x(\rho)$ *différentiable en* $\rho = 0$ ;
  $x(\underline{f_0})$ *et* $x(\overline{f_0})$ *différentiables en* $f_0 = f^*$ ;
- *[barrière]* $x(\rho)$ *minimiseur de* $\phi(x, \rho)$ ;
- *[centre]* $x(\overline{f_0})$ *minimiseur de* $\overline{\phi}(x, p, \overline{f_0})$ ;
- *[Karmarkar]* $p > q$ *assure que* $x(\underline{f_0})$ *minimiseur de* $\underline{\phi}(x, p, \underline{f_0})$.

# Paramétrisations de la trajectoire centrale

## Hypothèses

- Indépendance linéaire des contraintes actives (LICQ) ;
- complémentarité stricte (CS) ; $|\{i : \lambda_i > 0\}| = q$ ;
- conditions suffisantes de second ordre en $x^*$ (SOSC).

## Théorème

- *[trajectoire]* $x(\rho)$ *différentiable en* $\rho = 0$ ;
  $x(\underline{f_0})$ *et* $x(\overline{f_0})$ *différentiables en* $f_0 = f^*$ ;
- *[barrière]* $x(\rho)$ *minimiseur de* $\phi(x, \rho)$ ;
- *[centre]* $x(\overline{f_0})$ *minimiseur de* $\overline{\phi}(x, p, \overline{f_0})$ ;
- *[Karmarkar]* $p > q$ *assure que* $x(\underline{f_0})$ *minimiseur de* $\underline{\phi}(x, p, \underline{f_0})$.

# Paramétrisations de la trajectoire centrale

## Hypothèses

- Indépendance linéaire des contraintes actives (LICQ) ;
- complémentarité stricte (CS) ; $|\{i : \lambda_i > 0\}| = q$ ;
- conditions suffisantes de second ordre en $x^*$ (SOSC).

## Théorème

- *[trajectoire]* $x(\rho)$ *différentiable en* $\rho = 0$ ;
  $x(\underline{f_0})$ *et* $x(\overline{f_0})$ *différentiables en* $f_0 = f^*$ ;
- *[barrière]* $x(\rho)$ *minimiseur de* $\phi(x, \rho)$ ;
- *[centre]* $x(\overline{f_0})$ *minimiseur de* $\overline{\phi}(x, p, \overline{f_0})$ ;
- *[Karmarkar]* $p > q$ *assure que* $x(\underline{f_0})$ *minimiseur de* $\underline{\phi}(x, p, \underline{f_0})$.

# Paramétrisations de la trajectoire centrale

## Hypothèses

- Indépendance linéaire des contraintes actives (LICQ) ;
- complémentarité stricte (CS) ; $|\{i : \lambda_i > 0\}| = q$ ;
- conditions suffisantes de second ordre en $x^*$ (SOSC).

## Théorème

- *[trajectoire]* $x(\rho)$ *différentiable en* $\rho = 0$ ;
  $x(\underline{f_0})$ *et* $x(\overline{f_0})$ *différentiables en* $f_0 = f^*$ ;
- *[barrière]* $x(\rho)$ *minimiseur de* $\phi(x, \rho)$ ;
- *[centre]* $x(\overline{f_0})$ *minimiseur de* $\overline{\phi}(x, p, \overline{f_0})$ ;
- *[Karmarkar]* $p > q$ *assure que* $x(\underline{f_0})$ *minimiseur de* $\underline{\phi}(x, p, \underline{f_0})$.

# Paramétrisations de la trajectoire centrale

## Hypothèses

- Indépendance linéaire des contraintes actives (LICQ) ;
- complémentarité stricte (CS) ; $|\{i : \lambda_i > 0\}| = q$ ;
- conditions suffisantes de second ordre en $x^*$ (SOSC).

## Théorème

- *[trajectoire]* $x(\rho)$ *différentiable en* $\rho = 0$ ; $x(\underline{f_0})$ *et* $x(\overline{f_0})$ *différentiables en* $f_0 = f^*$ ;
- *[barrière]* $x(\rho)$ *minimiseur de* $\phi(x, \rho)$ ;
- *[centre]* $x(\overline{f_0})$ *minimiseur de* $\overline{\phi}(x, p, \overline{f_0})$ ;
- *[Karmarkar]* $p > q$ *assure que* $x(\underline{f_0})$ *minimiseur de* $\underline{\phi}(x, p, \underline{f_0})$.

# Paramétrisations de la trajectoire centrale

## Hypothèses

- Indépendance linéaire des contraintes actives (LICQ) ;
- complémentarité stricte (CS) ; $|\{i : \lambda_i > 0\}| = q$ ;
- conditions suffisantes de second ordre en $x^*$ (SOSC).

## Théorème

- *[trajectoire]* $x(\rho)$ *différentiable en* $\rho = 0$ ;
  $x(\underline{f_0})$ *et* $x(\overline{f_0})$ *différentiables en* $f_0 = f^*$ ;
- *[barrière]* $x(\rho)$ *minimiseur de* $\phi(x, \rho)$ ;
- *[centre]* $x(\overline{f_0})$ *minimiseur de* $\overline{\phi}(x, p, \overline{f_0})$ ;
- *[Karmarkar]* $p > q$ *assure que* $x(\underline{f_0})$ *minimiseur de* $\underline{\phi}(x, p, \underline{f_0})$.

# Paramétrisations de la trajectoire centrale

## Hypothèses

- Indépendance linéaire des contraintes actives (LICQ) ;
- complémentarité stricte (CS) ; $|\{i : \lambda_i > 0\}| = q$ ;
- conditions suffisantes de second ordre en $x^*$ (SOSC).

## Théorème

- *[trajectoire]* $x(\rho)$ *différentiable en* $\rho = 0$ *;* $x(\underline{f_0})$ *et* $x(\overline{f_0})$ *différentiables en* $f_0 = f^*$ *;*
- *[barrière]* $x(\rho)$ *minimiseur de* $\phi(x, \rho)$ *;*
- *[centre]* $x(\overline{f_0})$ *minimiseur de* $\overline{\phi}(x, p, \overline{f_0})$ *;*
- *[Karmarkar]* $p > q$ *assure que* $x(\underline{f_0})$ *minimiseur de* $\underline{\phi}(x, p, \underline{f_0})$.

# Paramétrisations de la trajectoire centrale

## Hypothèses

- Indépendance linéaire des contraintes actives (LICQ) ;
- complémentarité stricte (CS) ; $|\{i : \lambda_i > 0\}| = q$ ;
- conditions suffisantes de second ordre en $x^*$ (SOSC).

## Théorème

- *[trajectoire]* $x(\rho)$ *différentiable en* $\rho = 0$ ; $x(\underline{f_0})$ *et* $x(\overline{f_0})$ *différentiables en* $f_0 = f^*$ ;
- *[barrière]* $x(\rho)$ *minimiseur de* $\phi(x, \rho)$ ;
- *[centre]* $x(\overline{f_0})$ *minimiseur de* $\overline{\phi}(x, p, \overline{f_0})$ ;
- *[Karmarkar]* $p > q$ *assure que* $x(\underline{f_0})$ *minimiseur de* $\underline{\phi}(x, p, \underline{f_0})$.

# Discussion des hypothèses

- LICQ semble trop forte : en programmation linéaire, on en a pas besoin ; des exemples et résultats préliminaires suggèrent que même MFCQ n'est pas nécessaire ;
- CS est nécessaire : Mifflin : $x(\rho) - x^* \sim O(\sqrt{\rho})$ sinon ;
- SOSC semble nécessaire : exemple de Gilbert, Gonzaga, Karas. avex $f$ convexe $C^\infty$.

Puisque les 3 paramétrisations semblent équivalentes, raisonnons sur $x(\rho)$.
$f^*$ inconnue, la paramétrisation en $\rho$ est plus facile à exploiter.

# Discussion des hypothèses

- LICQ semble trop forte : en programmation linéaire, on en a pas besoin ; des exemples et résultats préliminaires suggèrent que même MFCQ n'est pas nécessaire ;
- CS est nécessaire : Mifflin : $x(\rho) - x^* \sim O(\sqrt{\rho})$ sinon ;
- SOSC semble nécessaire : exemple de Gilbert, Gonzaga, Karas. avex $f$ convexe $C^\infty$.

Puisque les 3 paramétrisations semblent équivalentes, raisonnons sur $x(\rho)$.

$f^*$ inconnue, la paramétrisation en $\rho$ est plus facile à exploiter.

# Discussion des hypothèses

- LICQ semble trop forte : en programmation linéaire, on en a pas besoin ; des exemples et résultats préliminaires suggèrent que même MFCQ n'est pas nécessaire ;
- CS est nécessaire : Mifflin : $x(\rho) - x^* \sim O(\sqrt{\rho})$ sinon ;
- SOSC semble nécessaire : exemple de Gilbert, Gonzaga, Karas. avex $f$ convexe $C^\infty$.

Puisque les 3 paramétrisations semblent équivalentes, raisonnons sur $x(\rho)$.

$f^*$ inconnue, la paramétrisation en $\rho$ est plus facile à exploiter.

# Discussion des hypothèses

- LICQ semble trop forte : en programmation linéaire, on en a pas besoin ; des exemples et résultats préliminaires suggèrent que même MFCQ n'est pas nécessaire ;
- CS est nécessaire : Mifflin : $x(\rho) - x^* \sim O(\sqrt{\rho})$ sinon ;
- SOSC semble nécessaire : exemple de Gilbert, Gonzaga, Karas. avex $f$ convexe $C^\infty$.

Puisque les 3 paramétrisations semblent équivalentes, raisonnons sur $x(\rho)$.
$f^*$ inconnue, la paramétrisation en $\rho$ est plus facile à exploiter.

# Convergence linéaire — centres

- $x(r) - x^* \sim O(\rho)$. Donc, comme $\rho$ s'approche de 0 comparablement à $f_0$ qui s'approche de $f^*$, l'erreur converge linéairement en fonction de ces paramétrisations.
- Principe : considérer une suite $f_k$ ou $\rho_k$, résoudre approximativement un sous-problème pour la valeur $f_k$ ou $r_k$, puis passer à $k+1$.
- Approximativement : $\|\nabla\phi(x_k, f(x_{k-1}))\| \le \epsilon\rho$.

**Théorème (Mifflin 1976)**

$$\frac{(f(x_k) - f^*)}{(f(x_{k-1}) - f^*)} \longrightarrow \frac{\rho q}{1 + \rho q}$$

- Pas d'évaluation de l'effort requis pour obtenir $x_k$. Résultat semblable si LICQ n'est pas satisfaite !

# Convergence linéaire — centres

- $x(r) - x^* \sim O(\rho)$. Donc, comme $\rho$ s'approche de 0 comparablement à $f_0$ qui s'approche de $f^*$, l'erreur converge linéairement en fonction de ces paramétrisations.
- Principe : considérer une suite $f_k$ ou $\rho_k$, résoudre approximativement un sous-problème pour la valeur $f_k$ ou $r_k$, puis passer à $k+1$.
- Approximativement : $\|\nabla\phi(x_k, f(x_{k-1}))\| \leq \epsilon\rho$.

**Théorème (Mifflin 1976)**

$$\frac{(f(x_k) - f^*)}{(f(x_{k-1}) - f^*)} \longrightarrow \frac{\rho q}{1 + \rho q}$$

- Pas d'évaluation de l'effort requis pour obtenir $x_k$. Résultat semblable si LICQ n'est pas satisfaite !

# Convergence linéaire — centres

- $x(r) - x^* \sim O(\rho)$. Donc, comme $\rho$ s'approche de 0 comparablement à $f_0$ qui s'approche de $f^*$, l'erreur converge linéairement en fonction de ces paramétrisations.
- Principe : considérer une suite $f_k$ ou $\rho_k$, résoudre approximativement un sous-problème pour la valeur $f_k$ ou $r_k$, puis passer à $k+1$.
- Approximativement : $\|\nabla\phi(x_k, f(x_{k-1}))\| \leq \epsilon\rho$.

**Théorème (Mifflin 1976)**

$$\frac{(f(x_k) - f^*)}{(f(x_{k-1}) - f^*)} \longrightarrow \frac{\rho q}{1+\rho q}$$

- Pas d'évaluation de l'effort requis pour obtenir $x_k$. Résultat semblable si LICQ n'est pas satisfaite !

# Convergence linéaire — centres

- $x(r) - x^* \sim O(\rho)$. Donc, comme $\rho$ s'approche de 0 comparablement à $f_0$ qui s'approche de $f^*$, l'erreur converge linéairement en fonction de ces paramétrisations.
- Principe : considérer une suite $f_k$ ou $\rho_k$, résoudre approximativement un sous-problème pour la valeur $f_k$ ou $r_k$, puis passer à $k+1$.
- Approximativement : $\|\nabla\phi(x_k, f(x_{k-1}))\| \leq \epsilon\rho$.

**Théorème (Mifflin 1976)**

$\frac{(f(x_k)-f^*)}{(f(x_{k-1})-f^*)} \longrightarrow \frac{\rho q}{1+\rho q}$

- Pas d'évaluation de l'effort requis pour obtenir $x_k$. Résultat semblable si LICQ n'est pas satisfaite !

UNIVERSITÉ DE SHERBROOKE

# Convergence linéaire — centres

- $x(r) - x^* \sim O(\rho)$. Donc, comme $\rho$ s'approche de 0 comparablement à $f_0$ qui s'approche de $f^*$, l'erreur converge linéairement en fonction de ces paramétrisations.
- Principe : considérer une suite $f_k$ ou $\rho_k$, résoudre approximativement un sous-problème pour la valeur $f_k$ ou $r_k$, puis passer à $k+1$.
- Approximativement : $\|\nabla\phi(x_k, f(x_{k-1}))\| \leq \epsilon\rho$.

**Théorème (Mifflin 1976)**

$$\frac{(f(x_k)-f^*)}{(f(x_{k-1})-f^*)} \longrightarrow \frac{\rho q}{1+\rho q}$$

- Pas d'évaluation de l'effort requis pour obtenir $x_k$. Résultat semblable si LICQ n'est pas satisfaite !

# Convergence linéaire—barrière

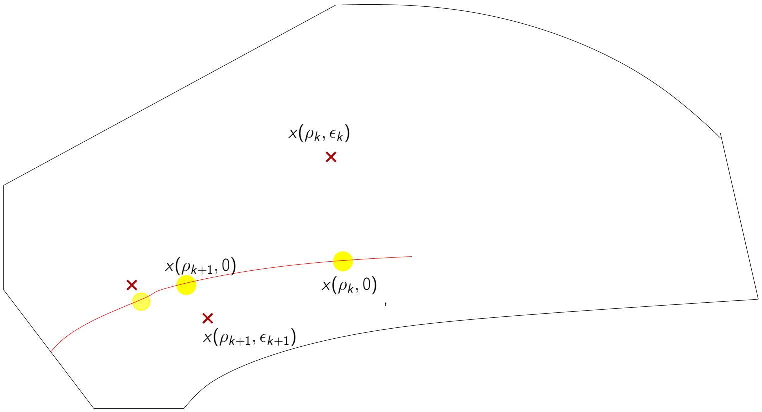
Lorsque $\rho_k \searrow 0$, mauvais conditionnement.
Approximativement : $\|\nabla\phi(x_k, \rho_k)\| \leq \rho_k$.

**Lemme**

$\|\nabla\phi(x_k, \rho_{k+1})\| \sim O(\frac{\rho_k}{\rho_{k+1}})$.

$\frac{\rho_k}{\rho_{k+1}} < \infty$ empêche la convergence plus rapide que linéaire de la suite $\rho_k$.

# Itérations illustrées



# Extrapolation — centre

- Jarre, Sonnevend et Stoer (1988) proposent d'utiliser $x_k, x_{k-1}, \ldots x_{k-l}$ pour construire une fonction d'extrapolation estimant $x_{k+1}$.
- Pas d'analyse d'ordre de convergence.
- Exemples numériques encourageants.
- Difficulté : on espère utiliser une tolérance beaucoup plus grande pour $x_{k-i}$ que pour $x_k$, qualité de l'extrapolation limitée par ces grandes tolérances.

# Extrapolation — centre

- Jarre, Sonnevend et Stoer (1988) proposent d'utiliser $x_k, x_{k-1}, \ldots x_{k-l}$ pour construire une fonction d'extrapolation estimant $x_{k+1}$.
- Pas d'analyse d'ordre de convergence.
- Exemples numériques encourageants.
- Difficulté : on espère utiliser une tolérance beaucoup plus grande pour $x_{k-i}$ que pour $x_k$, qualité de l'extrapolation limitée par ces grandes tolérances.

# Extrapolation — centre

- Jarre, Sonnevend et Stoer (1988) proposent d'utiliser $x_k, x_{k-1}, \ldots x_{k-l}$ pour construire une fonction d'extrapolation estimant $x_{k+1}$.
- Pas d'analyse d'ordre de convergence.
- Exemples numériques encourageants.
- Difficulté : on espère utiliser une tolérance beaucoup plus grande pour $x_{k-i}$ que pour $x_k$, qualité de l'extrapolation limitée par ces grandes tolérances.

# Extrapolation — centre

- Jarre, Sonnevend et Stoer (1988) proposent d'utiliser $x_k, x_{k-1}, \ldots x_{k-l}$ pour construire une fonction d'extrapolation estimant $x_{k+1}$.
- Pas d'analyse d'ordre de convergence.
- Exemples numériques encourageants.
- Difficulté : on espère utiliser une tolérance beaucoup plus grande pour $x_{k-i}$ que pour $x_k$, qualité de l'extrapolation limitée par ces grandes tolérances.

# Extrapolation — barrière

- Équations primales-duales perturbées :

$$\Phi(x, \lambda, t, \rho) = \begin{cases} -\nabla f(x) + \sum \lambda_i \nabla g_i(x) & = & t\bar{\epsilon} = \varepsilon \\ \lambda_i g_i(x) - \rho & = & 0, \end{cases}$$

- $\bar{\epsilon} = \epsilon/\|\epsilon\|$ et $\|\epsilon\| = t \leq \rho$.
- $x(\rho, t)$ trajectoire différentiable en $\rho$ et $t$.
- $\hat{x}^r$ obtenu par une extrapolation d'ordre $r$ de $(\rho, t)$ à $(\rho^+, 0)$.
- $\|\hat{x}^r - x(\rho^+, 0)\| \sim O(\max(t, \rho)^{r+1})$.
- Principe
  - prédicteur : $\hat{x}$
  - correcteur : itérations de Newton pour obtenir $x^+$.

# Extrapolation — barrière

- Équations primales-duales perturbées :

$$\Phi(x, \lambda, t, \rho) = \begin{cases} -\nabla f(x) + \sum \lambda_i \nabla g_i(x) & = & t\bar{\epsilon} = \varepsilon \\ \lambda_i g_i(x) - \rho & = & 0, \end{cases}$$

- $\bar{\epsilon} = \epsilon / \|\epsilon\|$ et $\|\epsilon\| = t \leq \rho$.
- $x(\rho, t)$ trajectoire différentiable en $\rho$ et $t$.
- $\hat{x}^r$ obtenu par une extrapolation d'ordre $r$ de $(\rho, t)$ à $(\rho^+, 0)$.
- $\|\hat{x}^r - x(\rho^+, 0)\| \sim O(\max(t, \rho)^{r+1})$.
- Principe
  - prédicteur : $\hat{x}$
  - correcteur : itérations de Newton pour obtenir $x^+$.

# Extrapolation — barrière

- Équations primales-duales perturbées :

$$\Phi(x, \lambda, t, \rho) = \begin{cases} -\nabla f(x) + \sum \lambda_i \nabla g_i(x) & = & t\bar{\epsilon} = \varepsilon \\ \lambda_i g_i(x) - \rho & = & 0, \end{cases}$$

- $\bar{\epsilon} = \epsilon / \|\epsilon\|$ et $\|\epsilon\| = t \leq \rho$.
- $x(\rho, t)$ trajectoire différentiable en $\rho$ et $t$.
- $\hat{x}^r$ obtenu par une extrapolation d'ordre $r$ de $(\rho, t)$ à $(\rho^+, 0)$.
- $\|\hat{x}^r - x(\rho^+, 0)\| \sim O(\max(t, \rho)^{r+1})$.
- Principe
  - prédicteur : $\hat{x}$
  - correcteur : itérations de Newton pour obtenir $x^+$.

# Extrapolation — barrière

- Équations primales-duales perturbées :

$$\Phi(x, \lambda, t, \rho) = \begin{cases} -\nabla f(x) + \sum \lambda_i \nabla g_i(x) & = & t\bar{\epsilon} = \varepsilon \\ \lambda_i g_i(x) - \rho & = & 0, \end{cases}$$

- $\bar{\epsilon} = \epsilon / \|\epsilon\|$ et $\|\epsilon\| = t \leq \rho$.
- $x(\rho, t)$ trajectoire différentiable en $\rho$ et $t$.
- $\hat{x}^r$ obtenu par une extrapolation d'ordre $r$ de $(\rho, t)$ à $(\rho^+, 0)$.
- $\|\hat{x}^r - x(\rho^+, 0)\| \sim O(\max(t, \rho)^{r+1})$.
- Principe
  - prédicteur : $\hat{x}$
  - correcteur : itérations de Newton pour obtenir $x^+$.

# Extrapolation — barrière

- Équations primales-duales perturbées :

$$\Phi(x,\lambda,t,\rho) = \begin{cases} -\nabla f(x) + \sum \lambda_i \nabla g_i(x) & = & t\bar{\epsilon} = \varepsilon \\ \lambda_i g_i(x) - \rho & = & 0, \end{cases}$$

- $\bar{\epsilon} = \epsilon / \|\epsilon\|$ et $\|\epsilon\| = t \leq \rho$.
- $x(\rho, t)$ trajectoire différentiable en $\rho$ et $t$.
- $\hat{x}^r$ obtenu par une extrapolation d'ordre $r$ de $(\rho, t)$ à $(\rho^+, 0)$.
- $\|\hat{x}^r - x(\rho^+, 0)\| \sim O(\max(t,\rho)^{r+1})$.
- Principe
  - prédicteur : $\hat{x}$
  - correcteur : itérations de Newton pour obtenir $x^+$.

# Extrapolation — barrière

- Équations primales-duales perturbées :

$$\Phi(x, \lambda, t, \rho) = \begin{cases} -\nabla f(x) + \sum \lambda_i \nabla g_i(x) & = & t\bar{\epsilon} = \varepsilon \\ \lambda_i g_i(x) - \rho & = & 0, \end{cases}$$

- $\bar{\epsilon} = \epsilon / \|\epsilon\|$ et $\|\epsilon\| = t \leq \rho$.
- $x(\rho, t)$ trajectoire différentiable en $\rho$ et $t$.
- $\hat{x}^r$ obtenu par une extrapolation d'ordre $r$ de $(\rho, t)$ à $(\rho^+, 0)$.
- $\|\hat{x}^r - x(\rho^+, 0)\| \sim O(\max(t, \rho)^{r+1})$.
- Principe
  - prédicteur : $\hat{x}$
  - correcteur : itérations de Newton pour obtenir $x^+$.

# Extrapolation — barrière

- Équations primales-duales perturbées :

$$\Phi(x, \lambda, t, \rho) = \begin{cases} -\nabla f(x) + \sum \lambda_i \nabla g_i(x) & = & t\bar{\epsilon} = \varepsilon \\ \lambda_i g_i(x) - \rho & = & 0, \end{cases}$$

- $\bar{\epsilon} = \epsilon / \|\epsilon\|$ et $\|\epsilon\| = t \leq \rho$.
- $x(\rho, t)$ trajectoire différentiable en $\rho$ et $t$.
- $\hat{x}^r$ obtenu par une extrapolation d'ordre $r$ de $(\rho, t)$ à $(\rho^+, 0)$.
- $\|\hat{x}^r - x(\rho^+, 0)\| \sim O(\max(t, \rho)^{r+1})$.
- Principe
  - prédicteur : $\hat{x}$
  - correcteur : itérations de Newton pour obtenir $x^+$.

# Extrapolation — barrière

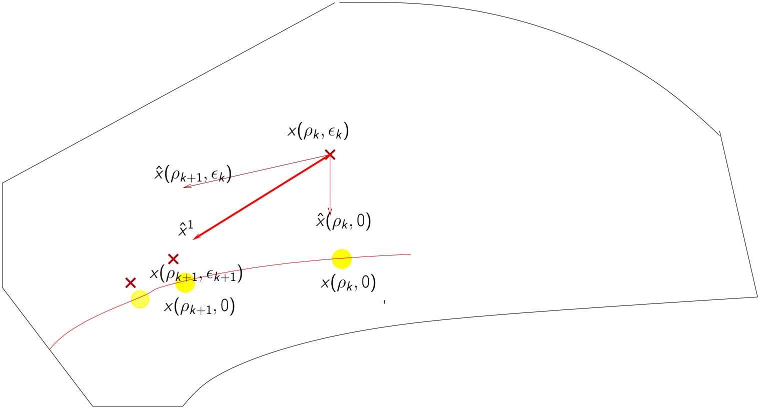
- Équations primales-duales perturbées :

$$\Phi(x, \lambda, t, \rho) = \begin{cases} -\nabla f(x) + \sum \lambda_i \nabla g_i(x) & = & t\bar{\epsilon} = \varepsilon \\ \lambda_i g_i(x) - \rho & = & 0, \end{cases}$$

- $\bar{\epsilon} = \epsilon / \|\epsilon\|$ et $\|\epsilon\| = t \leq \rho$.
- $x(\rho, t)$ trajectoire différentiable en $\rho$ et $t$.
- $\hat{x}^r$ obtenu par une extrapolation d'ordre $r$ de $(\rho, t)$ à $(\rho^+, 0)$.
- $\|\hat{x}^r - x(\rho^+, 0)\| \sim O(\max(t, \rho)^{r+1})$.
- Principe
    - prédicteur : $\hat{x}$
    - correcteur : itérations de Newton pour obtenir $x^+$.

UNIVERSITÉ DE SHERBROOKE

# Extrapolation illustrées



# Lemmes généraux

## Lemme (1)

*Soit $\hat{x}^r$ tel que $\|\hat{x}^r - x(\rho^+, 0)\| \sim \rho^{r+1}$. Alors,*
$\nabla\phi(\hat{x}^r, \rho^+) \sim \mathcal{O}(\frac{\rho^{r+1}}{\rho^+}))$.

## Lemme

*Soit $\hat{x}^1$ tel que $\|\hat{x}^1 - x(\rho^+, 0)\| \sim \rho^2$. Alors,*
$d_N = -\nabla^2\phi(\hat{x}^1, \rho^+)^{-1}\nabla\phi(\hat{x}^1, \rho^+)^t \sim \mathcal{O}(\rho^2))$.

## Lemme

*Soit $\hat{x}^1$ tel que $\|\hat{x}^1 - x(\rho^+, 0)\| \sim \rho^2$. Alors,*
$\nabla\phi(\hat{x}^1 + d_N, \rho^+) \sim \mathcal{O}(\frac{\rho^4}{\rho^{+2}})$.

# Lemmes généraux

## Lemme (1)

*Soit* $\hat{x}^r$ *tel que* $\|\hat{x}^r - x(\rho^+, 0)\| \sim \rho^{r+1}$. *Alors,*
$\nabla \phi(\hat{x}^r, \rho^+) \sim \mathcal{O}(\frac{\rho^{r+1}}{\rho^+}))$.

## Lemme

*Soit* $\hat{x}^1$ *tel que* $\|\hat{x}^1 - x(\rho^+, 0)\| \sim \rho^2$. *Alors,*
$d_N = -\nabla^2 \phi(\hat{x}^1, \rho^+)^{-1} \nabla \phi(\hat{x}^1, \rho^+)^t \sim \mathcal{O}(\rho^2))$.

## Lemme

*Soit* $\hat{x}^1$ *tel que* $\|\hat{x}^1 - x(\rho^+, 0)\| \sim \rho^2$. *Alors,*
$\nabla \phi(\hat{x}^1 + d_N, \rho^+) \sim \mathcal{O}(\frac{\rho^4}{\rho^{+2}})$.

# Lemmes généraux

## Lemme (1)

*Soit* $\hat{x}^r$ *tel que* $\|\hat{x}^r - x(\rho^+, 0)\| \sim \rho^{r+1}$. *Alors,*
$\nabla\phi(\hat{x}^r, \rho^+) \sim \mathcal{O}(\frac{\rho^{r+1}}{\rho^+}))$.

## Lemme

*Soit* $\hat{x}^1$ *tel que* $\|\hat{x}^1 - x(\rho^+, 0)\| \sim \rho^2$. *Alors,*
$d_N = -\nabla^2\phi(\hat{x}^1, \rho^+)^{-1}\nabla\phi(\hat{x}^1, \rho^+)^t \sim \mathcal{O}(\rho^2))$.

## Lemme

*Soit* $\hat{x}^1$ *tel que* $\|\hat{x}^1 - x(\rho^+, 0)\| \sim \rho^2$. *Alors,*
$\nabla\phi(\hat{x}^1 + d_N, \rho^+) \sim \mathcal{O}(\frac{\rho^4}{\rho^{+2}})$.

# Extrapolation d'ordre 1

- $\hat{x}^1 = x(\rho, t) + \frac{\partial x}{\partial \rho}(\rho^+ - \rho) + \frac{\partial x}{\partial t}(-t)$

**Théorème**

*$\frac{\rho^+}{\rho^{4/3}} < \infty$ assure une seule correction de Newton ;*
*$\frac{\rho^+}{\rho^{8/5}} < \infty$ assure deux corrections de Newton et convergence asymptotique optimale (critère d'Ostrovski)*

- Coût d'extrapolation et de correction comparables : solution d'un système linéaire.
-
$$\hat{\omega}^1 = \begin{pmatrix} \hat{x}^1 \\ \hat{\lambda}^1 \end{pmatrix} = \nabla_{x\lambda}\Phi^{-1}\begin{pmatrix} -\varepsilon \\ (\rho^+ - \rho)e \end{pmatrix}$$

# Extrapolation d'ordre 1

- $\hat{x}^1 = x(\rho, t) + \frac{\partial x}{\partial \rho}(\rho^+ - \rho) + \frac{\partial x}{\partial t}(-t)$

**Théorème**

*$\frac{\rho^+}{\rho^{4/3}} < \infty$ assure une seule correction de Newton ;*
*$\frac{\rho^+}{\rho^{8/5}} < \infty$ assure deux corrections de Newton et convergence asymptotique optimale (critère d'Ostrovski)*

- Coût d'extrapolation et de correction comparables : solution d'un système linéaire.
-
$$\hat{\omega}^1 = \begin{pmatrix} \hat{x}^1 \\ \hat{\lambda}^1 \end{pmatrix} = \nabla_{x\lambda}\Phi^{-1}\begin{pmatrix} -\varepsilon \\ (\rho^+ - \rho)e \end{pmatrix}$$

# Extrapolation d'ordre 1

- $\hat{x}^1 = x(\rho, t) + \frac{\partial x}{\partial \rho}(\rho^+ - \rho) + \frac{\partial x}{\partial t}(-t)$

**Théorème**

*$\frac{\rho^+}{\rho^{4/3}} < \infty$ assure une seule correction de Newton ;*
*$\frac{\rho^+}{\rho^{8/5}} < \infty$ assure deux corrections de Newton et convergence asymptotique optimale (critère d'Ostrovski)*

- Coût d'extrapolation et de correction comparables : solution d'un système linéaire.
- $$\hat{\omega}^1 = \begin{pmatrix} \hat{x}^1 \\ \hat{\lambda}^1 \end{pmatrix} = \nabla_{x\lambda}\Phi^{-1}\begin{pmatrix} -\varepsilon \\ (\rho^+ - \rho)e \end{pmatrix}$$

# Extrapolation d'ordre 1

- $\hat{x}^1 = x(\rho, t) + \frac{\partial x}{\partial \rho}(\rho^+ - \rho) + \frac{\partial x}{\partial t}(-t)$

**Théorème**

$\frac{\rho^+}{\rho^{4/3}} < \infty$ *assure une seule correction de Newton ;*
$\frac{\rho^+}{\rho^{8/5}} < \infty$ *assure deux corrections de Newton et convergence asymptotique optimale (critère d'Ostrovski)*

- Coût d'extrapolation et de correction comparables : solution d'un système linéaire.
- $$\hat{\omega}^1 = \begin{pmatrix} \hat{x}^1 \\ \hat{\lambda}^1 \end{pmatrix} = \nabla_{x\lambda}\Phi^{-1}\begin{pmatrix} -\varepsilon \\ (\rho^+ - \rho)e \end{pmatrix}$$

# Extrapolation d'ordre p

- Avec un ordre d'extrapolation d'ordre supérieur, plus besoin de correction.
- $\hat{x}^2 = \hat{x}^1 + \frac{\partial^2 x}{\partial \rho^2}(\rho^+ - \rho)^2 + \frac{\partial^2 x}{\partial t^2}(-t)^2 + 2\frac{\partial^2 x}{\partial \rho \partial t}(\rho^+ - \rho)(-t)$

**Théorème**

*En utilisant l'extrapolation d'ordre r, $\hat{x}$ satisfait à*

$\|\nabla \phi(\hat{x}^r_{k+1}, \rho_{k+1})\| \leq \rho_{k+1}$ *dès que* $\frac{\rho_k^{\frac{r+1}{2}}}{\rho_{k+1}} \to 0.$

- L'ordre limite est donc linéaire lorsque $r = 1$, $\frac{3}{2}$ pour $r = 2$, quadratique pour $r = 3$ et ainsi de suite.

# Extrapolation d'ordre p

- Avec un ordre d'extrapolation d'ordre supérieur, plus besoin de correction.
- $\hat{x}^2 = \hat{x}^1 + \frac{\partial^2 x}{\partial \rho^2}(\rho^+ - \rho)^2 + \frac{\partial^2 x}{\partial t^2}(-t)^2 + 2\frac{\partial^2 x}{\partial \rho \partial t}(\rho^+ - \rho)(-t)$

**Théorème**

*En utilisant l'extrapolation d'ordre r, $\hat{x}$ satisfait à*

$$\|\nabla\phi(\hat{x}^r_{k+1}, \rho_{k+1})\| \leq \rho_{k+1} \text{ dès que } \frac{\rho_k^{\frac{r+1}{2}}}{\rho_{k+1}} \to 0.$$

- L'ordre limite est donc linéaire lorsque $r = 1$, $\frac{3}{2}$ pour $r = 2$, quadratique pour $r = 3$ et ainsi de suite.

UNIVERSITÉ DE SHERBROOKE

# Extrapolation d'ordre p

- Avec un ordre d'extrapolation d'ordre supérieur, plus besoin de correction.
- $\hat{x}^2 = \hat{x}^1 + \frac{\partial^2 x}{\partial \rho^2}(\rho^+ - \rho)^2 + \frac{\partial^2 x}{\partial t^2}(-t)^2 + 2\frac{\partial^2 x}{\partial \rho \partial t}(\rho^+ - \rho)(-t)$

**Théorème**

*En utilisant l'extrapolation d'ordre r, $\hat{x}$ satisfait à*

$\|\nabla \phi(\hat{x}^r_{k+1}, \rho_{k+1})\| \leq \rho_{k+1}$ *dès que* $\frac{\rho_k^{\frac{r+1}{2}}}{\rho_{k+1}} \to 0.$

- L'ordre limite est donc linéaire lorsque $r = 1$, $\frac{3}{2}$ pour $r = 2$, quadratique pour $r = 3$ et ainsi de suite.

# Extrapolation d'ordre p

- Avec un ordre d'extrapolation d'ordre supérieur, plus besoin de correction.
- $\hat{x}^2 = \hat{x}^1 + \frac{\partial^2 x}{\partial \rho^2}(\rho^+ - \rho)^2 + \frac{\partial^2 x}{\partial t^2}(-t)^2 + 2\frac{\partial^2 x}{\partial \rho \partial t}(\rho^+ - \rho)(-t)$

**Théorème**

*En utilisant l'extrapolation d'ordre $r$, $\hat{x}$ satisfait à*

$\|\nabla \phi(\hat{x}^r_{k+1}, \rho_{k+1})\| \leq \rho_{k+1}$ *dès que* $\frac{\rho_k^{\frac{r+1}{2}}}{\rho_{k+1}} \to 0.$

- L'ordre limite est donc linéaire lorsque $r = 1$, $\frac{3}{2}$ pour $r = 2$, quadratique pour $r = 3$ et ainsi de suite.

# Extrapolation d'ordre p — coût de calcul

- Coût(obtenir $\nabla\Phi$) >> coût(obtenir $\Phi$).
- On peut récrire l'expression des $\ddot{\omega}$ :

$$\ddot{\omega}_{\theta_1\theta_2} = \nabla_{x\lambda}\Phi^{-1}\left(\nabla_{x\lambda}\left(\nabla_{x\lambda}(\Phi\dot{\omega}_{\theta_2})\dot{\omega}_{\theta_1}\right)\right)$$

- Même matrice que Newton et $\hat{\omega}^1$.
- Coût($\nabla_{x\lambda}\left(\nabla_{x\lambda}(\Phi\dot{\omega}_{\theta_2})\right)\dot{\omega}_{\theta_1}$) $\sim$ coût(obtenir $\Phi$).
- Conclusion : si on choisit Newton, autant continuer avec $\hat{\omega}^2$ !

Exemple :

| opération | $\Phi$ | $\nabla Phi$ | $LL^t = \nabla Phi$ | $LL\omega = b$ |
|---|---|---|---|---|
| coût | $mn^2$ | $mn^3$ | $(n+m)^3$ | $(n+m)^2)$ |

# Extrapolation d'ordre p — coût de calcul

- Coût(obtenir $\nabla\Phi$) $>>$ coût(obtenir $\Phi$).
- On peut récrire l'expression des $\ddot{\omega}$ :

$$\ddot{\omega}_{\theta_1\theta_2} = \nabla_{x\lambda}\Phi^{-1}\left(\nabla_{x\lambda}\left(\nabla_{x\lambda}(\Phi\dot{\omega}_{\theta_2})\dot{\omega}_{\theta_1}\right)\right)$$

- Même matrice que Newton et $\hat{\omega}^1$.
- Coût$\left(\nabla_{x\lambda}\left(\nabla_{x\lambda}(\Phi\dot{\omega}_{\theta_2})\right)\dot{\omega}_{\theta_1}\right) \sim$ coût(obtenir $\Phi$).
- Conclusion : si on choisit Newton, autant continuer avec $\hat{\omega}^2$ !

Exemple :

| opération | $\Phi$ | $\nabla Phi$ | $LL^t = \nabla Phi$ | $LL\omega = b$ |
|---|---|---|---|---|
| coût | $mn^2$ | $mn^3$ | $(n+m)^3$ | $(n+m)^2)$ |

UNIVERSITÉ DE SHERBROOKE

# Extrapolation d'ordre p — coût de calcul

- Coût(obtenir $\nabla\Phi$) $>>$ coût(obtenir $\Phi$).
- On peut récrire l'expression des $\ddot{\omega}$ :

$$\ddot{\omega}_{\theta_1\theta_2} = \nabla_{x\lambda}\Phi^{-1}\left(\nabla_{x\lambda}\left(\nabla_{x\lambda}(\Phi\dot{\omega}_{\theta_2})\dot{\omega}_{\theta_1}\right)\right)$$

- Même matrice que Newton et $\hat{\omega}^1$.
- Coût$\left(\nabla_{x\lambda}\left(\nabla_{x\lambda}(\Phi\dot{\omega}_{\theta_2})\right)\dot{\omega}_{\theta_1}\right) \sim$ coût(obtenir $\Phi$).
- Conclusion : si on choisit Newton, autant continuer avec $\hat{\omega}^2$ !

Exemple :

| opération | $\Phi$ | $\nabla Phi$ | $LL^t = \nabla Phi$ | $LL\omega = b$ |
|---|---|---|---|---|
| coût | $mn^2$ | $mn^3$ | $(n+m)^3$ | $(n+m)^2)$ |

# Extrapolation d'ordre p — coût de calcul

- Coût(obtenir $\nabla\Phi$) $>>$ coût(obtenir $\Phi$).
- On peut récrire l'expression des $\ddot{\omega}$ :

$$\ddot{\omega}_{\theta_1\theta_2} = \nabla_{x\lambda}\Phi^{-1}\left(\nabla_{x\lambda}\left(\nabla_{x\lambda}(\Phi\dot{\omega}_{\theta_2})\dot{\omega}_{\theta_1}\right)\right)$$

- Même matrice que Newton et $\hat{\omega}^1$.
- Coût$\left(\nabla_{x\lambda}\left(\nabla_{x\lambda}(\Phi\dot{\omega}_{\theta_2})\right)\dot{\omega}_{\theta_1}\right) \sim$ coût(obtenir $\Phi$).
- Conclusion : si on choisit Newton, autant continuer avec $\hat{\omega}^2$ !

Exemple :

| opération | $\Phi$ | $\nabla Phi$ | $LL^t = \nabla Phi$ | $LL\omega = b$ |
|---|---|---|---|---|
| coût | $mn^2$ | $mn^3$ | $(n+m)^3$ | $(n+m)^2)$ |

# Extrapolation d'ordre p — coût de calcul

- Coût(obtenir $\nabla\Phi$) $>>$ coût(obtenir $\Phi$).
- On peut récrire l'expression des $\ddot{\omega}$ :

$$\ddot{\omega}_{\theta_1\theta_2} = \nabla_{x\lambda}\Phi^{-1}\left(\nabla_{x\lambda}\left(\nabla_{x\lambda}(\Phi\dot{\omega}_{\theta_2})\dot{\omega}_{\theta_1}\right)\right)$$

- Même matrice que Newton et $\hat{\omega}^1$.
- Coût$\left(\nabla_{x\lambda}\left(\nabla_{x\lambda}(\Phi\dot{\omega}_{\theta_2})\right)\dot{\omega}_{\theta_1}\right) \sim$ coût(obtenir $\Phi$).
- Conclusion : si on choisit Newton, autant continuer avec $\hat{\omega}^2$ !

Exemple :

| opération | $\Phi$ | $\nabla Phi$ | $LL^t = \nabla Phi$ | $LL\omega = b$ |
|---|---|---|---|---|
| coût | $mn^2$ | $mn^3$ | $(n+m)^3$ | $(n+m)^2)$ |

UNIVERSITÉ DE SHERBROOKE

# Extrapolation d'ordre p — coût de calcul

- Coût(obtenir $\nabla\Phi$) $>>$ coût(obtenir $\Phi$).
- On peut récrire l'expression des $\ddot{\omega}$ :

$$\ddot{\omega}_{\theta_1\theta_2} = \nabla_{x\lambda}\Phi^{-1}\left(\nabla_{x\lambda}\left(\nabla_{x\lambda}(\Phi\dot{\omega}_{\theta_2})\dot{\omega}_{\theta_1}\right)\right)$$

- Même matrice que Newton et $\hat{\omega}^1$.
- Coût$\left(\nabla_{x\lambda}\left(\nabla_{x\lambda}(\Phi\dot{\omega}_{\theta_2})\right)\dot{\omega}_{\theta_1}\right) \sim$ coût(obtenir $\Phi$).
- Conclusion : si on choisit Newton, autant continuer avec $\hat{\omega}^2$ !

Exemple :

| opération | $\Phi$ | $\nabla Phi$ | $LL^t = \nabla Phi$ | $LL\omega = b$ |
|---|---|---|---|---|
| coût | $mn^2$ | $mn^3$ | $(n+m)^3$ | $(n+m)^2)$ |

# Minimisations d'ordre 1

Éviter d'évaluer et utiliser $\nabla\Phi$

- Les stratégies d'ordre 1 deviennent intéressantes.
- Gradient conjugué, L-BFGS, BFGS... : convergence superlinéaire ?
- Extrapolations "sécantes" ?
- Résolution itérative de $\nabla\Phi\omega = b$ : convergence $O(\rho^{1+\gamma})$ ?.

# Minimisations d'ordre 1

Éviter d'évaluer et utiliser $\nabla\Phi$

- Les stratégies d'ordre 1 deviennent intéressantes.
- Gradient conjugué, L-BFGS, BFGS... : convergence superlinéaire ?
- Extrapolations "sécantes" ?
- Résolution itérative de $\nabla\Phi\omega = b$ : convergence $O(\rho^{1+\gamma})$ ?.

# Minimisations d'ordre 1

Éviter d'évaluer et utiliser $\nabla\Phi$

- Les stratégies d'ordre 1 deviennent intéressantes.
- Gradient conjugué, L-BFGS, BFGS... : convergence superlinéaire ?
- Extrapolations "sécantes" ?
- Résolution itérative de $\nabla\Phi\omega = b$ : convergence $O(\rho^{1+\gamma})$ ?.

# Minimisations d'ordre 1

Éviter d'évaluer et utiliser $\nabla\Phi$

- Les stratégies d'ordre 1 deviennent intéressantes.
- Gradient conjugué, L-BFGS, BFGS... : convergence superlinéaire ?
- Extrapolations "sécantes" ?
- Résolution itérative de $\nabla\Phi\omega = b$ : convergence $O(\rho^{1+\gamma})$ ?.

# Plan

# Résumé

- Méthode des centres, de barrière, Karmarkar unifiés.
- Paramétrisation barrière plus facile car on connaît la cible $\rho = 0$.
- Prédicteur d'ordre supérieur semble plus efficace.
- Sans utiliser $\nabla\Phi$, réévaluer les prédiction utilisant d'anciens itérés et les paramétrisations "centre" et "Karmarkar".

# Résumé

- Méthode des centres, de barrière, Karmarkar unifiés.
- Paramétrisation barrière plus facile car on connaît la cible $\rho = 0$.
- Prédicteur d'ordre supérieur semble plus efficace.
- Sans utiliser $\nabla\Phi$, réévaluer les prédiction utilisant d'anciens itérés et les paramétrisations "centre" et "Karmarkar".

UNIVERSITÉ DE SHERBROOKE

# Résumé

- Méthode des centres, de barrière, Karmarkar unifiés.
- Paramétrisation barrière plus facile car on connaît la cible $\rho = 0$.
- Prédicteur d'ordre supérieur semble plus efficace.
- Sans utiliser $\nabla\Phi$, réévaluer les prédiction utilisant d'anciens itérés et les paramétrisations "centre" et "Karmarkar".

UNIVERSITÉ DE SHERBROOKE

# Résumé

- Méthode des centres, de barrière, Karmarkar unifiés.
- Paramétrisation barrière plus facile car on connaît la cible $\rho = 0$.
- Prédicteur d'ordre supérieur semble plus efficace.
- Sans utiliser $\nabla\Phi$, réévaluer les prédiction utilisant d'anciens itérés et les paramétrisations "centre" et "Karmarkar".

# Omissions

Méthodes primales–duales.

- Linéaire, convexe : méthodes de choix, convergence quadratique.
- Non-linéaire : Wächter&Biegler : ATTENTION !
- Convergence prédicteur seulement, sans correction $O(\rho^{1+\gamma})$, $\gamma < 1$ (Gould, Orban, Sartenauer, Toint 2000)

## Omissions

Méthodes primales–duales.

- Linéaire, convexe : méthodes de choix, convergence quadratique.
- Non-linéaire : Wächter&Biegler : ATTENTION !
- Convergence prédicteur seulement, sans correction $O(\rho^{1+\gamma})$, $\gamma < 1$ (Gould, Orban, Sartenauer, Toint 2000)

# Omissions

Méthodes primales–duales.

- Linéaire, convexe : méthodes de choix, convergence quadratique.
- Non-linéaire : Wächter&Biegler : ATTENTION !
- Convergence prédicteur seulement, sans correction $O(\rho^{1+\gamma})$, $\gamma < 1$ (Gould, Orban, Sartenauer, Toint 2000)

# Lectures I

Fiacco&McCormick
*Nonlinear programming : Sequential Unconstrained Minimization Techniques.*
SIAM, 1990.

Pierre Huard
La méthode des centres.