Quelques mots sur la forme révisée de l'algorithme (primal) simplexe

Anthony Przybylski, Xavier Gandibleux

Université de Nantes, M1 ORO

Pédagogie vs réalité

- Présentation pédagogique de l'algorithme du simplexe
  - Utilisation du tableau simplexe comme structure de données
  - Utilisation d'une itération de la méthode du pivot de Gauss pour mettre à jour le tableau
- Inconvénients
  - Propagation très rapide d'imprécisions numériques dans le tableau simplexe
  - Coût temporel du maintien du tableau
  - Aucune exploitation du caractère creux de la matrice des contraintes
- Implémentations
  - Abandon du tableau simplexe
  - Utilisation/adaptation de notions d'analyse numérique
  - Toujours des progrès en combinant des expérimentations massives (nombreux paramètres dans la méthode) et des éléments théoriques

Pédagogie vs réalité

- Présentation pédagogique de l'algorithme du simplexe
  - Utilisation du tableau simplexe comme structure de données
  - Utilisation d'une itération de la méthode du pivot de Gauss pour mettre à jour le tableau
- Inconvénients
  - Propagation très rapide d'imprécisions numériques dans le tableau simplexe
  - Coût temporel du maintien du tableau
  - Aucune exploitation du caractère creux de la matrice des contraintes
- Implémentations
  - Abandon du tableau simplexe
  - Utilisation/adaptation de notions d'analyse numérique
  - Toujours des progrès en combinant des expérimentations massives (nombreux paramètres dans la méthode) et des éléments théoriques

Pédagogie vs réalité

- Présentation pédagogique de l'algorithme du simplexe
  - Utilisation du tableau simplexe comme structure de données
  - Utilisation d'une itération de la méthode du pivot de Gauss pour mettre à jour le tableau
- Inconvénients
  - Propagation très rapide d'imprécisions numériques dans le tableau simplexe
  - Coût temporel du maintien du tableau
  - Aucune exploitation du caractère creux de la matrice des contraintes
- Implémentations
  - Abandon du tableau simplexe
  - Utilisation/adaptation de notions d'analyse numérique
  - Toujours des progrès en combinant des expérimentations massives (nombreux paramètres dans la méthode) et des éléments théoriques

Algorithme du simplexe : point de vue algébrique (rappel)

Un PL sous forme standard peut s'écrire

$$\begin{array}{rcl} \max\ z & = & c^T x \\ s.c. \quad Ax & = & b \\ x & \geq & 0 \end{array}$$

- Soit $\mathcal{B}$ une base, la sous-matrice $A_{\mathcal{B}}$ définie par les colonnes de A qui sont dans $\mathcal{B}$ est donc inversible
- On note alors $\mathcal{N}$ l'ensemble des variables hors-bases, la sous-matrice de A définie par les colonnes de A qui sont dans $\mathcal{N}$ est notée $A_{\mathcal{N}}$
- La séparation entre les variables de base et hors-bases peut aussi se faire dans les vecteurs c ($c_{\mathcal{B}}$ et $c_{\mathcal{N}}$) et x ($x_{\mathcal{B}}$ et $x_{\mathcal{N}}$)

Algorithme du simplexe : point de vue algébrique (rappel)

- $A = (A_{\mathcal{B}}, A_{\mathcal{N}})$, $c^T = (c_{\mathcal{B}}^T, c_{\mathcal{N}}^T)$ and $x = (x_{\mathcal{B}}^T, x_{\mathcal{N}}^T)^T$
- $$\begin{aligned} Ax = b &\iff (A_{\mathcal{B}}, A_{\mathcal{N}})(x_{\mathcal{B}}^T, x_{\mathcal{N}}^T)^T = b \\ &\iff A_{\mathcal{B}}x_{\mathcal{B}} + A_{\mathcal{N}}x_{\mathcal{N}} = b \\ &\iff A_{\mathcal{B}}x_{\mathcal{B}} = b - A_{\mathcal{N}}x_{\mathcal{N}} \\ &\iff x_{\mathcal{B}} = A_{\mathcal{B}}^{-1}(b - A_{\mathcal{N}}x_{\mathcal{N}}) \end{aligned}$$
  En fixant $x_{\mathcal{N}} = 0$, on obtient $x_{\mathcal{B}} = A_{\mathcal{B}}^{-1}b$. La solution de base associée à $\mathcal{B}$ est donc $(x_{\mathcal{B}}, 0)$
- $$\begin{aligned} c^T x &= (c_{\mathcal{B}}^T, c_{\mathcal{N}}^T)(x_{\mathcal{B}}^T, x_{\mathcal{N}}^T)^T \\ &= c_{\mathcal{B}}^T x_{\mathcal{B}} + c_{\mathcal{N}}^T x_{\mathcal{N}} \\ &= c_{\mathcal{B}}^T A_{\mathcal{B}}^{-1} b + (c_{\mathcal{N}}^T - c_{\mathcal{B}}^T A_{\mathcal{B}}^{-1} A_{\mathcal{N}}) x_{\mathcal{N}} \end{aligned}$$
  Vecteur des coûts réduits : $\bar{c}^T = c^T - c_{\mathcal{B}}^T A_{\mathcal{B}}^{-1} A$
  $\bar{c}^T = (\bar{c}_{\mathcal{B}}^T, \bar{c}_{\mathcal{N}})^T$ avec $\bar{c}_{\mathcal{B}} = 0$

Algorithme du simplexe : point de vue algébrique (rappel)

- $A = (A_{\mathcal{B}}, A_{\mathcal{N}})$, $c^T = (c_{\mathcal{B}}^T, c_{\mathcal{N}}^T)$ and $x = (x_{\mathcal{B}}^T, x_{\mathcal{N}}^T)^T$
- $$\begin{aligned} Ax = b &\iff (A_{\mathcal{B}}, A_{\mathcal{N}})(x_{\mathcal{B}}^T, x_{\mathcal{N}}^T)^T = b \\ &\iff A_{\mathcal{B}} x_{\mathcal{B}} + A_{\mathcal{N}} x_{\mathcal{N}} = b \\ &\iff A_{\mathcal{B}} x_{\mathcal{B}} = b - A_{\mathcal{N}} x_{\mathcal{N}} \\ &\iff x_{\mathcal{B}} = A_{\mathcal{B}}^{-1}(b - A_{\mathcal{N}} x_{\mathcal{N}}) \end{aligned}$$
  En fixant $x_{\mathcal{N}} = 0$, on obtient $x_{\mathcal{B}} = A_{\mathcal{B}}^{-1} b$. La solution de base associée à $\mathcal{B}$ est donc $(x_{\mathcal{B}}, 0)$
- $$\begin{aligned} c^T x &= (c_{\mathcal{B}}^T, c_{\mathcal{N}}^T)(x_{\mathcal{B}}^T, x_{\mathcal{N}}^T)^T \\ &= c_{\mathcal{B}}^T x_{\mathcal{B}} + c_{\mathcal{N}}^T x_{\mathcal{N}} \\ &= c_{\mathcal{B}}^T A_{\mathcal{B}}^{-1} b + (c_{\mathcal{N}}^T - c_{\mathcal{B}}^T A_{\mathcal{B}}^{-1} A_{\mathcal{N}}) x_{\mathcal{N}} \end{aligned}$$
  Vecteur des coûts réduits : $\bar{c}^T = c^T - c_{\mathcal{B}}^T A_{\mathcal{B}}^{-1} A$
  $\bar{c}^T = (\bar{c}_{\mathcal{B}}^T, \bar{c}_{\mathcal{N}})^T$ avec $\bar{c}_{\mathcal{B}} = 0$

Algorithme du simplexe : point de vue algébrique (rappel)

- $A = (A_{\mathcal{B}}, A_{\mathcal{N}})$, $c^T = (c_{\mathcal{B}}^T, c_{\mathcal{N}}^T)$ and $x = (x_{\mathcal{B}}^T, x_{\mathcal{N}}^T)^T$
- $$\begin{aligned} Ax = b &\iff (A_{\mathcal{B}}, A_{\mathcal{N}})(x_{\mathcal{B}}^T, x_{\mathcal{N}}^T)^T = b \\ &\iff A_{\mathcal{B}}x_{\mathcal{B}} + A_{\mathcal{N}}x_{\mathcal{N}} = b \\ &\iff A_{\mathcal{B}}x_{\mathcal{B}} = b - A_{\mathcal{N}}x_{\mathcal{N}} \\ &\iff x_{\mathcal{B}} = A_{\mathcal{B}}^{-1}(b - A_{\mathcal{N}}x_{\mathcal{N}}) \end{aligned}$$
  En fixant $x_{\mathcal{N}} = 0$, on obtient $x_{\mathcal{B}} = A_{\mathcal{B}}^{-1}b$. La solution de base associée à $\mathcal{B}$ est donc $(x_{\mathcal{B}}, 0)$
- $$\begin{aligned} c^Tx &= (c_{\mathcal{B}}^T, c_{\mathcal{N}}^T)(x_{\mathcal{B}}^T, x_{\mathcal{N}}^T)^T \\ &= c_{\mathcal{B}}^T x_{\mathcal{B}} + c_{\mathcal{N}}^T x_{\mathcal{N}} \\ &= c_{\mathcal{B}}^T A_{\mathcal{B}}^{-1} b + (c_{\mathcal{N}}^T - c_{\mathcal{B}}^T A_{\mathcal{B}}^{-1} A_{\mathcal{N}}) x_{\mathcal{N}} \end{aligned}$$
  Vecteur des coûts réduits : $\bar{c}^T = c^T - c_{\mathcal{B}}^T A_{\mathcal{B}}^{-1} A$
  $\bar{c}^T = (\bar{c}_{\mathcal{B}}^T, \bar{c}_{\mathcal{N}})^T$ avec $\bar{c}_{\mathcal{B}} = 0$

Algorithme du simplexe : point de vue algébrique (rappel)

Tableau simplexe associé à la base $\mathcal{B}$

$\bar{c}$	$c^T - c_{\mathcal{B}}^T A_{\mathcal{B}}^{-1} A$	$-c_{\mathcal{B}}^T A_{\mathcal{B}}^{-1} b$
$x_{\mathcal{B}}$	$A_{\mathcal{B}}^{-1} A$	$A_{\mathcal{B}}^{-1} b$

Conséquences

- Il est possible de reconstruire le tableau simplexe associé à chaque base (complètement ou partiellement) à partir du tableau initial/des données du problème
- Utiliser ces formules réduit considérablement les imprécisions numériques mais pour un coût temporel prohibitif
- En particulier, même si $A_{\mathcal{B}}$ est creuse, $A_{\mathcal{B}}^{-1}$ ne l'est pas forcément

La méthode du simplexe en forme révisée

Grandes lignes

- À chaque itération, on sait quelles variables sont en base, ainsi que les valeurs des variables de base mais les éléments du tableau simplexe ne sont pas connus
- Nécessité de calculer les/des coûts réduits pour choisir une variable entrante
- Nécessité de calculer les éléments de la colonne de la variable entrante pour choisir une variable sortante, et mettre à jour les valeurs des variables de base

Calcul des coûts réduits

- Les coûts réduits des variables de base sont nuls
- Les coûts réduits des variables hors-bases sont donnés par

$$\bar{c}_{\mathcal{N}}^T = c_{\mathcal{N}}^T - c_{\mathcal{B}}^T A_{\mathcal{B}}^{-1} A_{\mathcal{N}}$$

- On résout $y^T = c_{\mathcal{B}}^T A_{\mathcal{B}}^{-1} \iff y^T A_{\mathcal{B}} = c_{\mathcal{B}}^T$
  où y est un vecteur d'inconnues, on a donc un système d'équations à résoudre
- On utilise y dans $c_{\mathcal{N}}^T - y^T A_{\mathcal{N}}$ pour obtenir les coûts réduits des variables hors-bases
- Remarque : il n'est pas forcément nécessaire de calculer tous les coûts réduits pour choisir une variable entrante

Exemple (1/4)

On considère la seconde itération de la résolution du PL suivant

$$
\begin{array}{rcrrrrrcl}
\max\ z & = & 3x_1 & +2x_2 & & & & & \\
 & s.c. & 2x_1 & +x_2 & +x_3 & & & = & 6 \\
 & & x_1 & +x_2 & & +x_4 & & = & 4 \\
 & & & x_2 & & & +x_5 & = & 3 \\
 & & x_1, & x_2, & x_3, & x_4, & x_5 & \geq & 0
\end{array}
$$

On a $\mathcal{B} = \{x_1, x_4, x_5\}$, $x_{\mathcal{B}} = (3, 1, 3)$ et $A_{\mathcal{B}} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Exemple (2/4)

On résout

$$y^T A_{\mathcal{B}} = c_{\mathcal{B}}^T \iff (y_1, y_2, y_3) \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = (3, 0, 0)$$

soit

$$\left\{ \begin{array}{rrrcl} 2y_1 & +y_2 & & = & 3 \\ & y_2 & & = & 0 \\ & & y_3 & = & 0 \end{array} \right.$$

On obtient $y^T = (\frac{3}{2}, 0, 0)$ qu'on injecte dans $\bar{c}_{\mathcal{N}}^T = c_{\mathcal{N}}^T - y^T A_{\mathcal{N}}$

$$(\bar{c}_2, \bar{c}_3) = (2, 0) - \left(\frac{3}{2}, 0, 0\right) \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} = \left(\frac{1}{2}, -\frac{3}{2}\right)$$

Calcul de la colonne de la variable entrante

- Les éléments de la colonne de la variable entrante (dans le tableau simplexe) sont donnés par

$$A_{\mathcal{B}}^{-1}a$$

  où a est la colonne de A dont l'indice correspond à la variable entrante
- On résout $d = A_{\mathcal{B}}^{-1}a \Longleftrightarrow A_{\mathcal{B}}d = a$
  où d est un vecteur d'inconnues, on a donc un système d'équations à résoudre

Exemple (3/4)

On choisit x_2 comme variable entrante, on a donc $a = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

$$A_{\mathcal{B}} d = a \iff \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

soit

$$\left\{ \begin{array}{cccccc} 2d_1 & & & = & 1 \\ d_1 & +d_2 & & = & 1 \\ & & d_3 & = & 1 \end{array} \right.$$

On obtient $d = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$

Choix de la variable sortante et MAJ des valeurs des variables de base

- On connaît
  - les valeurs des variables de base (habituelle colonne de droite du tableau simplexe)
  - la colonne de la variable entrante
- On peut donc
  - choisir la variable sortante
  - mettre à jour les variables de base

  de la façon habituelle

- On a $d = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$ et $x_{\mathcal{B}} = \begin{pmatrix} x_1 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 3 \end{pmatrix}$

  On obtient donc les ratios $\begin{pmatrix} \frac{3}{\frac{1}{2}} \\ \frac{1}{\frac{1}{2}} \\ \frac{3}{1} \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ 3 \end{pmatrix}$

  La variable sortante est donc x_4
- $\mathcal{B}$ devient $(\{x_1, x_4, x_5\} \cup \{x_2\}) \setminus \{x_4\} = \{x_1, x_2, x_5\}$

  Le vecteur des variables de base devient
  $\begin{pmatrix} x_1 \\ x_2 \\ x_5 \end{pmatrix} = \begin{pmatrix} 3 - \frac{1}{2} \times 2 \\ 2 \\ 3 - 1 \times 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$

Et le changement de base?

- Dans la version pédagogique de l'algorithme du simplexe
  - Toutes les informations du tableau simplexe sont disponibles
  - Le changement de base demande de mettre à jour le tableau simplexe
- Dans la forme révisée de l'algorithme du simplexe
  - Les informations utiles du tableau simplexe doivent être calculées
  - Il n'y aucun traitement à faire (dans cette version basique) pour le changement de base à part la mise-à-jour des valeurs des variables de base

Récapitulatif

Une itération de l'algorithme du simplexe en forme révisée (cas de maximisation)

Entrée: base (primale) admissible $\mathcal{B}$ et solution associée $(x_{\mathcal{B}}, 0)$

- Résoudre le système d'équations $yA_{\mathcal{B}} = c_{\mathcal{B}}$
- Choix d'une variable entrante x_r vérifiant $(\bar{c}_r =) c_r - ya > 0$ où a est la colonne de $A_{\mathcal{N}}$ correspondant à x_r
- Résoudre le système d'équations $A_{\mathcal{B}} d = a$
- Choix d'une variable sortante x_s : première variable s'annulant avec l'augmentation de x_r
- $\mathcal{B} \leftarrow (\mathcal{B} \setminus \{r\}) \cup \{s\}$ et actualiser $x_{\mathcal{B}}$

Sortie: Base (primale) admissible $\mathcal{B}$ et solution associée $(x_{\mathcal{B}}, 0)$

Pour une implémentation vraiment efficace

- Le point essentiel est de résoudre efficacement les deux systèmes d'équations
- Pour cela, de nombreuses solutions issues/adaptées de l'algèbre linéaire et l'analyse numérique sont possibles
- En particulier, de nombreuses méthodes ont été proposées pour résoudre des systèmes d'équations en exploitant le caractère creux du système
- Après un changement de base, seule une colonne de $A_{\mathcal{B}}$ est changée $\Rightarrow$ inutile de résoudre intégralement les systèmes d'équations
- Idée principale : décomposer la matrice $A_{\mathcal{B}}$ et maintenir cette décomposition en cas de changement de base
- Les solveurs modernes ont recours à une décomposition LU

Décomposition LU

- Une décomposition de $A_{\mathcal{B}} = LU$, où L est une matrice triangulaire inférieure, et U est une matrice triangulaire supérieure, est maintenue
- La résolution d'un système d'équations du type $A_{\mathcal{B}}d = a$ où d est le vecteur des inconnues devient $LUd = a$
- En posant $y = Ud$, on a d'abord le système triangulaire $Ly = a$ (immédiat) à résoudre
- Ensuite, le système triangulaire $Ud = y$ peut être directement résolu
- Conclusion : chaque système d'équations à résoudre dans la méthode du simplexe révisé demande juste de résoudre deux systèmes d'équations triangulaires

Exemple (1/7)

On considère la seconde itération de la résolution du PL suivant

$$\begin{array}{rrrrrrrr}
\max\ z = & 3x_1 & +2x_2 & & & & & \\
s.c. & 2x_1 & +x_2 & +x_3 & & & = & 6 \\
 & x_1 & +x_2 & & +x_4 & & = & 4 \\
 & & x_2 & & & +x_5 & = & 3 \\
 & x_1, & x_2, & x_3, & x_4, & x_5 & \geq & 0
\end{array}$$

On a $\mathcal{B} = \{x_1, x_4, x_5\}$, $x_{\mathcal{B}} = (3, 1, 3)$, $A_{\mathcal{B}} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$,

$L = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et $U = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

Exemple (2/7)

On veut résoudre $y^T A_{\mathcal{B}} = c_{\mathcal{B}}^T \iff y^T LU = c_{\mathcal{B}}^T$
Pour cela, on pose $t^T = y^T L$ et on résout

$$t^T U = c_{\mathcal{B}}^T \iff (t_1, t_2, t_3) \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = (3, 0, 0)$$

On obtient $t = (\frac{3}{2}, 0, 0)$ et on résout

$$y^T L = t^T \iff (y_1, y_2, y_3) \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} = (\frac{3}{2}, 0, 0)$$

On obtient $y^T = (\frac{3}{2}, 0, 0)$ qu'on injecte dans $\bar{c}_{\mathcal{N}}^T = c_{\mathcal{N}}^T - y^T A_{\mathcal{N}}$

$$(\bar{c}_2, \bar{c}_3) = (2, 0) - \left(\frac{3}{2}, 0, 0\right) \begin{pmatrix} 1 & 1 \\ 1 & 0 \\ 1 & 0 \end{pmatrix} = \left(\frac{1}{2}, -\frac{3}{2}\right)$$

Exemple (3/7)

On choisit x_2 comme variable entrante, on a donc $a = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$

On veut résoudre $A_{\mathcal{B}} d = a \Longleftrightarrow LUd = a$
Pour cela, on pose $y = Ud$ et on résout

$$Ly = a \Longleftrightarrow \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} y_1 \\ y_2 \\ y_3 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$$

On obtient $y = \begin{pmatrix} 1 \\ \frac{1}{2} \\ 1 \end{pmatrix}$ et on résout

$$Ud = y \Longleftrightarrow \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} d_1 \\ d_2 \\ d_3 \end{pmatrix} = \begin{pmatrix} 1 \\ \frac{1}{2} \\ 1 \end{pmatrix}$$

On obtient $d = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$

Exemple (4/7)

- On a $d = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{pmatrix}$ et $x_{\mathcal{B}} = \begin{pmatrix} x_1 \\ x_4 \\ x_5 \end{pmatrix} = \begin{pmatrix} 3 \\ 1 \\ 3 \end{pmatrix}$

  On obtient donc les ratios $\begin{pmatrix} \frac{3}{\frac{1}{2}} \\ \frac{1}{\frac{1}{2}} \\ \frac{3}{1} \end{pmatrix} = \begin{pmatrix} 6 \\ 2 \\ 3 \end{pmatrix}$

  La variable sortante est donc x_4
- $\mathcal{B}$ devient $\{x_1, x_4, x_5\} \cup \{x_2\} \setminus \{x_4\} = \{x_1, x_2, x_5\}$

  Le vecteur des variables de base devient

  $$\begin{pmatrix} x_1 \\ x_2 \\ x_5 \end{pmatrix} = \begin{pmatrix} 3 - \frac{1}{2} \times 2 \\ 2 \\ 3 - 1 \times 2 \end{pmatrix} = \begin{pmatrix} 2 \\ 2 \\ 1 \end{pmatrix}$$

Décomposition LU et changement de base (1/5)

- En plus de connaître les variables en base ainsi que leurs valeurs, une décomposition LU de la matrice $A_{\mathcal{B}}$ sera supposée connue au début de chaque itération
  $\Rightarrow$ Résolution rapide des deux systèmes d'équations
- Le changement de base demandera maintenant de mettre à jour la décomposition LU de la matrice $A_{\mathcal{B}}$ dont une colonne sera modifiée
- On note $\mathcal{B}$ la base avant changement, et $\mathcal{B}'$ la base après
  On connaît la décomposition $A_{\mathcal{B}} = LU$ et on doit obtenir $A_{\mathcal{B}'} = L'U'$

Décomposition LU et changement de base (2/5)

- $A_{\mathcal{B}} = LU \iff L^{-1}A_{\mathcal{B}} = U$
- $A_{\mathcal{B}}$ et $A_{\mathcal{B}'}$ ne diffèrent que d'une colonne
  (la colonne a dont l'indice dans A est le même que pour la variable entrante)
  $\Rightarrow$ Toutes les colonnes de $L^{-1}A_{\mathcal{B}'}$ sauf une sont directement issues de U
- $L^{-1}a$ a déjà été calculé en résolvant $Ly = a$
  (première moitié de la résolution de $A_{\mathcal{B}}d = a$)
- $L^{-1}A_{\mathcal{B}'}$ est donc obtenu sans calcul et est presque une matrice triangulaire supérieure

Décomposition LU et changement de base (3/5)

- Pour obtenir une matrice triangulaire supérieure, des éliminations gaussiennes sont appliquées
- On multiplie $L^{-1}A_{\mathcal{B}'}$ à gauche par des matrices du type

$$M_i = \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ & & & 1 & & \\ & & & m_i & \ddots & \\ & & & & & 1 \end{pmatrix}$$

où i correspond à la $i$ème colonne de la matrice
(Ce qui est équivalent à ajouter m_i fois la $i$ème ligne à la $(i+1)$ème)

Décomposition LU et changement de base (4/5)

- L'inverse M_i^{-1} d'une matrice Mi est donnée par

$$M_i^{-1} = \begin{pmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ & & & 1 & & \\ & & & -m_i & \ddots & \\ & & & & & 1 \end{pmatrix}$$

- De plus, la multiplication d'une matrice triangulaire inférieure à droite par des matrices du type M_i^{-1} préserve le caractère triangulaire inférieur

Décomposition LU et changement de base (5/5)

- On peut donc obtenir une matrice triangulaire supérieure U' donnée par

$$U' = M_l \dots M_k L^{-1} A_{\mathcal{B}'}$$

- Une matrice triangulaire inférieure L' est donnée par

$$L' = L M_k^{-1} \dots M_l^{-1}$$

- On a donc $L'U' = A_{\mathcal{B}'}$, la décomposition voulue est donc obtenue

Exemple (5/7)

- On a $\mathcal{B} = \{x_1, x_4, x_5\}$, $A_{\mathcal{B}} = \begin{pmatrix} 2 & 0 & 0 \\ 1 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$,

  $L = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$ et $U = \begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}$

- On a $\mathcal{B}' = \{x_1, x_2, x_5\}$, $A_{\mathcal{B}'} = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 1 & 0 \\ 0 & 1 & 1 \end{pmatrix}$

  et par la résolution de $Ly = a$, on a obtenu $y = \begin{pmatrix} 1 \\ \frac{1}{2} \\ 1 \end{pmatrix}$

  L', U'?

Exemple (6/7)

- On a immédiatement $L^{-1}A_{\mathcal{B}'} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 1 & 1 \end{pmatrix}$

  Seul un terme empêche la matrice d'être triangulaire supérieure

- Dans la méthode du pivot de Gauss, on ajouterait à la troisième ligne (-2) fois la seconde

  Ici, on multiplie à gauche la matrice $L^{-1}A_{\mathcal{B}'}$ par

$$M_2 = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -2 & 1 \end{pmatrix}$$

Exemple (7/7)

- On obtient $U' = M_2 L^{-1} A_{\mathcal{B}'} = \begin{pmatrix} 2 & 1 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & 1 \end{pmatrix}$
- L' est ensuite obtenu par $L' = L M_2^{-1}$

  On a $L' = \begin{pmatrix} 1 & 0 & 0 \\ \frac{1}{2} & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}$
- On a bien $L'U' = A_{\mathcal{B}'}$

Quelques remarques finales

- Il peut être utile de permuter des colonnes de $L^{-1}A_{\mathcal{B}'}$ avant de commencer les éliminations gaussiennes
- Des permutations de lignes sont aussi possibles (surtout pour des questions de précision numériques)
- La méthode présentée reste simple, de nombreux raffinements ont été proposés pour des raisons d'efficacité/de précision numérique
- Un exemple en petite taille n'est clairement pas convaincant
- Cependant, il ne faut pas oublier que les problèmes réels sont de grande taille, et que la matrice $A_{\mathcal{B}}$ est très creuse
- Les produits de matrices sont donc relativement rapides et stables numériquement

Quelques remarques finales

- Il peut être utile de permuter des colonnes de $L^{-1}A_{\mathcal{B}'}$ avant de commencer les éliminations gaussiennes
- Des permutations de lignes sont aussi possibles (surtout pour des questions de précision numériques)
- La méthode présentée reste simple, de nombreux raffinements ont été proposés pour des raisons d'efficacité/de précision numérique
- Un exemple en petite taille n'est clairement pas convaincant
- Cependant, il ne faut pas oublier que les problèmes réels sont de grande taille, et que la matrice $A_{\mathcal{B}}$ est très creuse
- Les produits de matrices sont donc relativement rapides et stables numériquement