Re-paramétrisation et réduction des systèmes irréductibles

Jean-Marc Cane[1], Arnaud Kubicki[1], Dominique Michelucci[1], Hichem Barki[2], Sebti Foufou[2]
[1] LE2I laboratory, Université de Bourgogne, BP 47870, 21078 Dijon France
[2] Dept. of Computer Science and Engineering, CENG, Qatar University, Doha, Qatar

24 mars 2014

Vous devez résoudre un système bien-contraint irréductible de taille $n = 1000$ équations et inconnues. Or Newton est en $O(n^3)$.

Le plus rageant : si vous connaissiez la valeur de $k \ll n$ **inconnues clefs (ou paramètres)**, alors le système serait réductible, et trivialement soluble en $O(n)$. Mais vous ne connaissez pas les valeurs de ces inconnues clefs.

Il est possible d'utiliser cette décomposition même quand les valeurs des inconnues clefs sont inconnues.

Re-paramétrisation et algèbre linéaire

De nombreux solveurs (Newton-Raphson, homotopie, Newton par intervalles) utilisent des procédures d'algèbre linéaire.

Donc exploiter la re-paramétrisation dans ces procédures semble une bonne idée : pas besoin de modifier les solveurs.

Ceci n'empêche d'ailleurs pas de tenter de profiter de la re-paramétrisation à un niveau plus élevé dans les algorithmes.

Plan

La re-paramétrisation, exemple 2D, exemple 3D

Rappels sur la théorie classique du couplage (à la base de la décomposition)

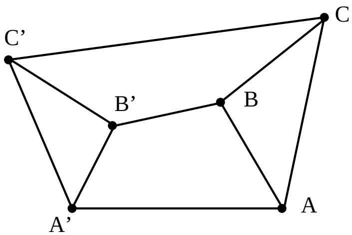
Mais elle ne réduit pas les systèmes irréductibles

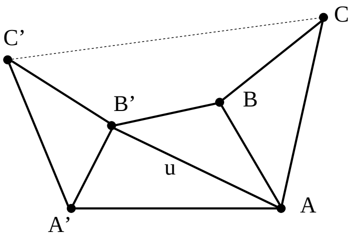
Comment tirer parti de la re-paramétrisation pour résoudre des systèmes linéaires

P-adique et remontée de Hensel, Solveurs par intervalles

Questions ouvertes, Conclusion

Re-paramétrisation, un exemple 2D





A droite : le système après re-paramétrisation. Si $u = AB'$ était connu, alors le système serait réductible et facile à résoudre ; l'équation fixant la longueur CC' serait redondante et pourrait être ignorée.

Le système d'équations (A', B', C' fixées)

$$0 = \text{dist}^2(A, B')^2 - u^2 \quad \text{définition de } u \tag{1}$$
$$0 = \text{dist}^2(A', A) - l^2_{A'A} \tag{2}$$
$$0 = \text{dist}^2(B', B) - l^2_{B'B} \tag{3}$$
$$0 = \text{dist}^2(A, B) - l^2_{AB} \tag{4}$$
$$0 = \text{dist}^2(A, C) - l^2_{AC} \tag{5}$$
$$0 = \text{dist}^2(B, C) - l^2_{BC} \tag{6}$$
$$0 = \text{dist}^2(C', C) - l^2_{C'C} \quad \text{équation oubliée} \tag{7}$$

où $l_{A'A}$, $l_{B'B}$, l_{AB}, l_{AC}, l_{BC}, $l_{C'C}$ sont les six longueurs spécifiées, et $\text{dist}^2(A, B)^2 = (x_A - x_B)^2 + (y_A - y_B)^2$.

Le jacobien a la structure suivante (une croix X représente une valeur non nulle) :

$$\begin{array}{|c|cccccc|}
\hline
u & x_A & y_A & x_B & y_B & x_C & y_C \\
\hline
\hline
X & X & X & 0 & 0 & 0 & 0 \\
0 & X & X & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & X & X & 0 & 0 \\
0 & X & X & X & X & 0 & 0 \\
0 & X & X & 0 & 0 & X & X \\
0 & 0 & 0 & X & X & X & X \\
\hline
0 & 0 & 0 & 0 & 0 & X & X \\
\hline
\end{array} \tag{8}$$

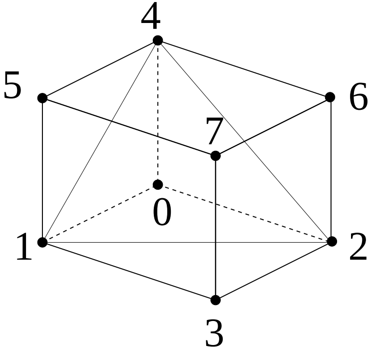
En entier, il est irréductible. Après suppression de la 1ère colonne et de la dernière ligne, il devient réductible : la matrice est triangulaire inférieure par blocs de 2x2.

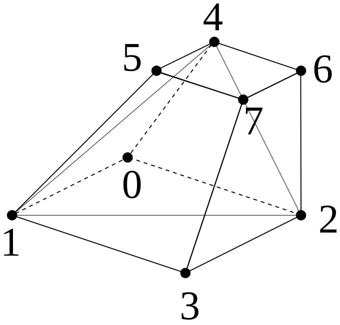
En fait, u est inutile. Il suffit d'utiliser x_A comme inconnue clef.

$$
\begin{array}{|c|cccccc|}
\hline
x_A & y_A & x_B & y_B & x_C & y_C \\
\hline
X & X & 0 & 0 & 0 & 0 \\
0 & 0 & X & X & 0 & 0 \\
X & X & X & X & 0 & 0 \\
X & X & 0 & 0 & X & X \\
0 & 0 & X & X & X & X \\
\hline
0 & 0 & 0 & 0 & X & X \\
\hline
\end{array}
\tag{9}
$$

Toujours une structure triangulaire inférieure par blocs bien visible.

Exemple 3D de re-paramétrisation : l'hexaèdre





Remarque : pour que le système soit bien contraint (modulo les déplacements), l'hexaèdre ne doit pas être plat.

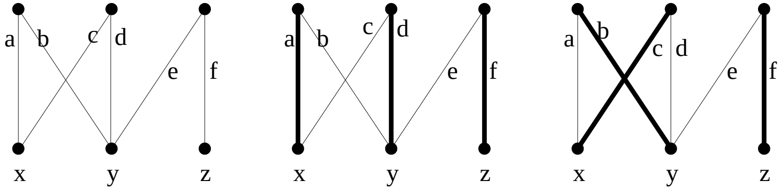
B-A-BA : la théorie classique du couplage

permet de décomposer des systèmes d'équations en 3 parties : sous-, bien-, sur-contrainte.

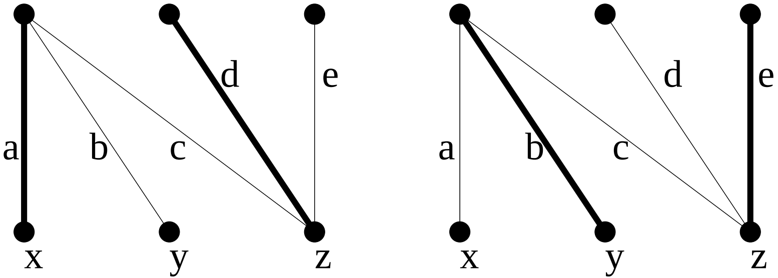
permet de décomposer un système bien-contraint en sous systèmes bien-contraint irréductibles.

mais elle ne peut rien faire pour les systèmes irréductibles.

Correspondance entre termes du jacobien et couplages

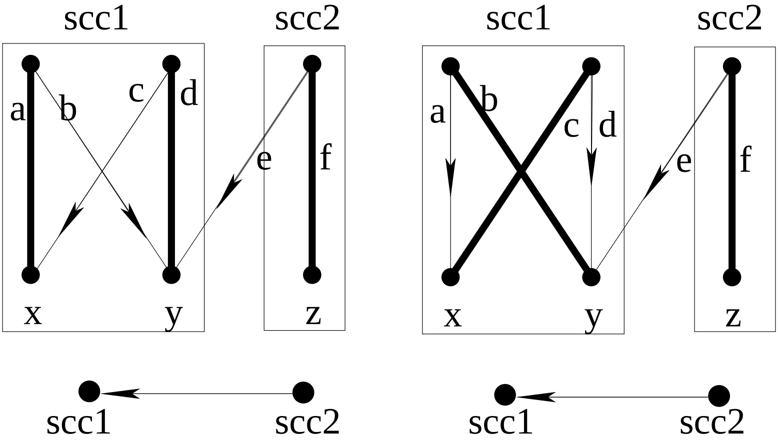


Le graphe biparti d'un système linéaire $ax + by = r_1, cx + dy = r_2, ey + fz = r_3$, et ses deux couplages parfaits adf et bcf. Le déterminant de la matrice sous jacente est $adf - bcf$.



Le système $F_1(x,y,z) = F_2(z) = F_3(z) = 0$ est structurellement mal contraint. "Structurellement" signifie que le système est mal contraint quelles que soient les valeurs de a,b,c,d,e. Aussi son graphe biparti n'a-t-il aucun couplage parfait.

Les couplages maximum contiennent deux arêtes (donc le rang de la matrice est 2, pour des valeurs génériques des coefficients a,b,c,d,e). Ce sont (a,d), (a,e), (b,d) et (b,e).


scc1
scc2
a
b
c
d
e
f
x
y
z
scc1
scc2
a
b
c
d
e
f
x
y
z
scc1
scc2
scc1
scc2

La décomposition en sous-systèmes accélère la résolution des systèmes linéaires ou non linéaires.

Calcul de l'inverse de M :

$$M = \begin{pmatrix} M_{1,1} & 0 & 0 \\ M_{2,1} & M_{2,2} & 0 \\ M_{3,1} & M_{3,2} & M_{3,3} \end{pmatrix} \tag{10}$$

Les blocs de $K = M^{-1}$ sont :

$$K_{l,c} = 0 \text{ quand } c > l \tag{11}$$

$$K_{l,l} = M_{l,l}^{-1} \tag{12}$$

$$K_{l,c} = -K_{l,l} \sum_{i=c}^{l-1} M_{l,i} K_{i,c} \text{ quand } l > c \tag{13}$$

En fait, il est toujours possible d'éviter l'inversion de la matrice M.

Pour résoudre $MX = B$:

$X_1 = \text{solve}(M_{1,1}X_1 = B_1)$,

$X_2 = \text{solve}(M_{2,2}X_2 = B_2 - M_{2,1}X_1)$,

$X_3 = \text{solve}(M_{3,3}X_3 = B_3 - (M_{3,1}X_1 + M_{3,2}X_2))$,

$X_k = \text{solve}(M_{k,k}X_k = B_k - \sum_{i=1}^{k-1} M_{k,i}X_i)$

M triangulaire inférieure : résoudre $MX = B$ est en $O(\alpha)$. Au pire, $\alpha = O(n^2)$: meilleur que $O(n^3)$. Souvent, $\alpha = C^{te}n$. En fait α est le nombre de termes non nuls dans M (TI).

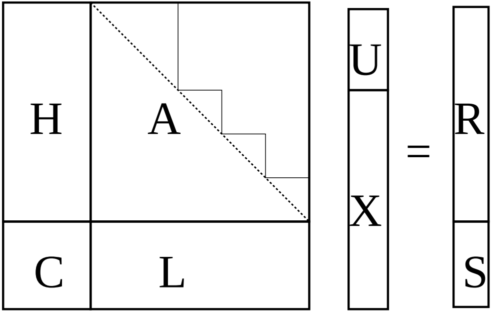
Etudions maintenant le cas re-paramétré

La méthode de Schreck, Imbach, Mathis, 2012 "Decomposition of geometric constraints with re-parameterization" donne les inconnues clefs (pas leurs valeurs !) et les équations redondantes.

Les systèmes sont bien-contraints et irréductibles.

Ils seraient réductibles si nous connaissions les valeurs des inconnues clefs.

Nous ne les connaissons pas.



$$HU + AX = R, CU + LX = S \tag{14}$$

A est carrée, inversible, triangulaire inférieure par bloc. C est k, k.

$$HU + AX = R, CU + LX = S \tag{15}$$

La méthode bête calcule A^{-1}, et comme $X = A^{-1}(R - HU)$:

$$CU + LX = S \Rightarrow CU + LA^{-1}(R - HU) = S \tag{16}$$
$$\Rightarrow \quad (C - LA^{-1}H)U = S - LA^{-1}R \tag{17}$$

$U = \text{solve}((C - LA^{-1}H)U = S - LA^{-1}R)$. $C - LA^{-1}H$ a une taille très petite devant celle de A. Pour l'hexaèdre, A est de taille 16×16, et C et $C - LA^{-1}H$ sont de taille 3×3. Ensuite $X = A^{-1}(R - HU)$.

$$HU + AX = R, CU + LX = S \tag{18}$$

Deuxième méthode, qui évite le calcul de M^{-1} :

A^{-1} apparaît d'abord dans $A^{-1}R$; posons $Z = A^{-1}R$. Alors $AZ = R$: un système linéaire avec matrice TIB, en $O(\alpha) \in O(n^2)$.

Ensuite A^{-1} apparaît dans $A^{-1}H$. Posons $Z = A^{-1}H$, donc $AZ = H$. Le nombre de colonnes de H, et de Z, est k ; donc $k \ll n$ systèmes linéaires.

Total : $(k+1)\alpha$ opérations, au pire kn^2, au lieu de n^3.

Ceci fonctionne quand les matrices H, A, C, U sont

des entiers (calcul modulaire ou P-adique)

des flottants : donc Newton-Raphson peut l'exploiter

des intervalles

Résoudre un système linéaire re-paramétré, avec A une matrice triangulaire inférieure par blocs est en

$$O(k\alpha + k^2 n + k^3)$$

où $O(\alpha)$ est le temps pour résoudre un système linéaire $Ax = b$.

α = le nombre de termes non nuls dans A quand A est TI.

$\alpha \in O(n^2)$: même dans le cas le pire, c'est beaucoup mieux que $O(n^3)$.

Le cas P-adique

Supposons que X_0 est une racine d'un système re-paramétré algébrique $F(X) = 0$ modulo p un nombre premier, ou une puissance d'un nombre premier.

Nous voulons calculer X_1 tel que $X_0 + pX_1$ soit une racine de $F(X) = 0$ modulo p^2. Taylor :

$$F(X_0 + pX_1) = F(X_0) + pF'(X_0)X_1 + \dots$$

Soit $\lambda = F(X_0)/p$. Alors $F(X_0 + pX_1) = 0$

$\Rightarrow F(X_0) + pF'(X_0)X_1 = 0 \; [\mod p^2] \Rightarrow \lambda + F'(X_0)X_1 = 0 \; [\mod p]$ un système linéaire, soluble modulo p.

Rappel : la remontée de Hensel n'est autre que la méthode de Newton, mais le résultat est exact. Attention : la re-paramétrisation n'aide pas pour trouver X_0....

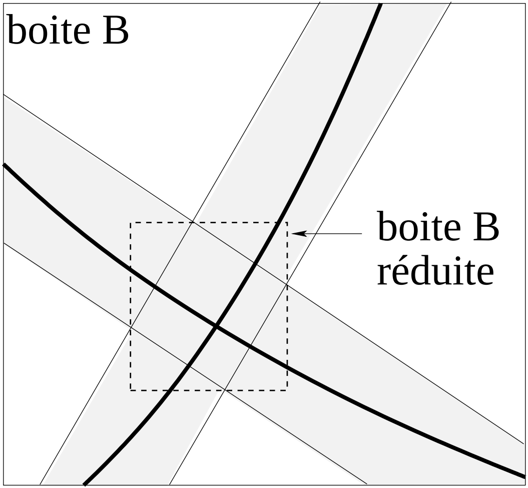
Re-paramétrisation et intervalles

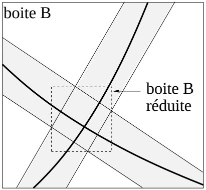
Le plus simple :

Résoudre par intervalles $f'(B)(x - x_0) = 0$, où x_0 est le centre de la boite étudiée B

Mais $f'(B)$ contient des intervalles, donc risque de singularité de certaines matrices sur la diagonale.

Solution : la forme centrée ; seul le membre droit contient des intervalles.


boite B
boite B
réduite



$$f'(B)(x - x_0) = 0 \Rightarrow \tag{19}$$
$$(J_0 + \Delta J)(x - x_0) = 0 \Rightarrow \tag{20}$$
$$J_0(x - x_0) + \Delta J(x - x_0) = 0 \Rightarrow \tag{21}$$
$$J_0(x - x_0) + \Delta J(B - x_0) = 0 \Rightarrow \tag{22}$$
$$J_0 \Delta x = -\Delta J(B - x_0) = V \Rightarrow \tag{23}$$
$$\Delta x = \text{solve}(J_0 \Delta x = V), x = x_0 + \Delta x \tag{24}$$

La méthode est compatible avec l'optimisation classique :
$x_{[k]} := (x_0 + \Delta x)_{[k]} \cap B_{[k]}$

Dès que l'intervalle $x_{[k]}$ est vide, il est prouvé que $B_{[k]}$ est vide, donc B est vide, donc B ne contient pas de racines.

Non compatible avec le pré-conditionnement, un des moyens de limiter l'effet enveloppant.

La re-paramétérisation à un plus haut niveau ?

Par exemple, pour un solveur par intervalles : ne couper les boites que selon les variables clefs. Les valeurs ([]) des inconnues non clefs sont calculées en fonction des inconnues clefs.

Attention : le solveur est modifié de fond en comble. Non testé !

Questions soulevées

- généralisation aux solveurs de Bernstein ?

- compatibilité avec pré-conditionnement ?

- pour trouver des ensembles clefs ? Méthode de Schreck, Imbach, Mathis Decomposition of geometric constraints with re parameterization, 2012.

- nous n'avons considéré que des systèmes d'équations, et pas de contraintes géométriques.

- la re-paramétrisation est elle utilisable ailleurs ? en algèbre linéaire (SVD, QR, etc), en programmation linéaire, en calcul formel... ?

Conclusion

Il est possible d'exploiter, du moins au niveau de l'algèbre linéaire, la décomposition qui serait possible si les valeurs des inconnues clefs étaient connues, même quand elles ne le sont pas !

Ceci permet de généraliser la méthode de Newton à de gros systèmes bien contraints irréductibles — si ils sont re-paramétrables avec $k \ll n$ inconnues clefs.

Il peut y avoir $k > 1$ inconnues clefs (dans l'article fondateur, Hoffmann, Gao, Yang : $k = 1$, tracaient la courbe $X(u)$, et détectait quand l'équation oubliée était satisfaite.)

Schreck, Mathis, Imbach ont proposé des méthodes efficaces pour re-paramétrer les systèmes de contraintes géométriques.