Techniques de l'Informatique, des Mathématiques, de la Microélectronique et de la Microscopie quantitative.
Unité associée au C.N.R.S. n° 397

UN ALGORITHME RAPIDE POUR LE CALCUL DE LA TRACE DE L'INVERSE D'UNE GRANDE MATRICE

D. GIRARD

RR 665 -M- Mai 1987

CENTRE NATIONAL DE LA RECHERCHE SCIENTIFIQUE
INSTITUT NATIONAL POLYTECHNIQUE DE GRENOBLE
UNIVERSITE SCIENTIFIQUE, TECHNOLOGIQUE ET MEDICALE DE GRENOBLE

Résumé: La trace d'une matrice symétrique A d'ordre n peut être estimée par la méthode de Monte-Carlo suivante: générer un vecteur w, pseudo n_échantillon de loi Gaussienne centrée de variance 1, calculer Aw, puis $s(w)= (w^tAw)/(w^tw)$. On montre que s(w) est un estimateur sans biais de $1/n$ tr(A), d'écart type relatif $\sqrt{2/(n+2)}$ fois la dispersion relative des valeurs propres. Cet algorithme peut diviser par n le coût des méthodes existantes pour le calcul de la trace de l'inverse d'une grande matrice creuse.

Abstract: An estimate of the trace of a nXn symmetric matrix A can be computed by the following Monte-Carlo algorithm: i)generate n pseudo-random values $w_1,...,w_n$, from the standard normal distribution and let $w=(w_1,...,w_n)$, ii)compute Aw, iii)calculate $s(w)= (w^tAw)/(w^tw)$ (and consider s(w) as an approximation to $1/n$ tr(A)). We show that this estimate is unbiased, and its standard deviation is $\sqrt{2/(n+2)}$ times the relative standard deviation of the eigenvalues of A. This algorithm may reduce by a factor n the cost of the actual methods for the trace of the inverse of a large sparse given matrix.

UN ALGORITHME RAPIDE POUR LE CALCUL DE LA TRACE DE L'INVERSE D'UNE GRANDE MATRICE

D. GIRARD

Mars 1987

1. Introduction

On propose et on examine ici un algorithme de type Monte-Carlo pour approcher la moyenne arithmétique des valeurs propres d'une matrice A d'ordre n (i.e. $^1/_n$ tr(A)), dont on ne connaît pas la diagonale, mais pour laquelle on sait calculer le vecteur Ay, pour tout y donné.

Actuellement la seule méthode généralement applicable consiste à calculer les n produits Ay pour $y=e_1,e_2,..,e_n$,les n vecteurs canoniques $e_1=(1,0,..,0),...$ et fournit donc en même temps toute la matrice A.

L' exemple principal que l'on a en tête est celui où A est donnée sous la forme $A = Z^{-1}$, l'inverse d'une grande matrice Z donnée pour laquelle on sait résoudre le système linéaire Zx=y avec un coût raisonnable (par exemple O(n)). Un tel problème apparaît dans le lissage par fonctions spline, ou dans le domaine de la restauration d'images, où y sont des données bruitées et Z un opérateur paramétré dont l'inversion produit les données lissées $x=Z^{-1}y$: ici la valeur de $tr(Z^{-1})$ fournit des informations statistiques importantes qui permettent d'évaluer si le lissage est bien adapté à y.

La méthode précédente demandant n résolutions, est très coûteuse quand n est grand. Signalons que, dans certains cas particuliers, les valeurs propres de Z peuvent être moins coûteuses à calculer (exemples des splines à pas équidistants [6], de la déconvolution [7] et de la tomographie [3]), et que, si Z est une matrice bande à 2m+1 diagonales, $O(^3/_2\, m^2n)$ opérations (et la place mémoire pour la décomposition de Cholesky de Z) suffisent pour calculer la diagonale de son inverse [2].

La méthode proposée ici est d'application générale, et consiste simplement à générer un vecteur w bruit blanc pseudo-aléatoire de variance 1, calculer Aw puis prendre comme estimateur de $^1/_n$ tr(A) le produit scalaire $g(w)= {}^1/_n\ w^t Aw$.

On étudie la précision que l'on peut attendre de cette méthode, qui repose sur le fait que l'écart type de cet estimateur vaut $\sqrt{^2/_n}$ fois la moyenne quadratique des valeurs propres.

On montre que la correction par la variance empirique de w qui conduit à l'estimateur $s(w)= {}^{w^t Aw}/_{w^t w}$ améliore encore cette précision.

Des techniques de réduction de variance sont étudiées et un exemple numérique montre finalement la puissance de cette méthode.

2. L'algorithme

Soit A une matrice n x n symétrique. Soit :

$$A = Q^t \ \mathrm{diag}(\lambda_1,\lambda_2,..,\lambda_n)\ Q$$

sa décomposition en valeurs propres où Q est unitaire et $\lambda_1,\lambda_2,..,\lambda_n$ sont les valeurs propres de A.

On notera E(X) l'espérance, $\sigma^2(X)=E(\ (X-E(X))^2\ )$ la variance de la variable aléatoire X.

Proposition 2.1.: Soient $\mu_1(A)$ la moyenne arithmétique des valeurs propres de A, $\mu_2(A)$ leur moyenne quadratique, et d(A) leur dispersion quadratique moyenne (dans la suite la désignation de la matrice fournissant ces valeurs sera omise si on parle de A), c'est-à-dire:

$$\mu_1(A)= \ {}^1\!/_n \sum_i \lambda_i$$

$$\mu_2(A)= \sqrt{{}^1\!/_n \sum_i \lambda_i^2}$$

$$d(A)= \sqrt{{}^1\!/_n \sum_i (\lambda_i - \mu_1)^2}$$

Soit w un vecteur de n variables aléatoires indépendantes, de loi Gaussienne centrée, de variance 1, (i.e. $w \sim> G(0,I)$).
Soit g(w) la variable aléatoire définie par

$$g(w)= {}^1\!/_n \ w^t A w\ .$$

Alors g(w) est un estimateur sans biais de μ_1 d'écart type $\sqrt{{}^2\!/_n}\ \mu_2$, c'est-à-dire:

$$E(g(w))={}^1\!/_n\ tr(A)=\mu_1$$

$$\sigma(g(w))= \sqrt{{}^2\!/_n}\ \ [\ {}^1\!/_n tr(A^2)\ ]^{1/2} = (\sqrt{{}^2\!/_n}\ )\ \mu_2.$$

démonst. : En notant que $z = Qw$ suit aussi la loi G(0,I) et que $g(w)= \ {}^1\!/_n\ \Sigma\ \lambda_i\ z_i^2$, la démonstration résulte des valeurs bien connues des moments:

$$E(z_i\ z_j\ ) =\delta_{i,j}$$

$$E(z_i^2 z_j^2) = \begin{cases} 3 \text{ si } i=j \\ 1 \text{ sinon,} \end{cases}$$

puisque:

$$E(g(w)) = 1/n \sum_i \lambda_i E(z_i^2),$$

et:

$$\begin{aligned} \sigma^2(g(w)) &= E(\, g(w)^2 - (E(g(w))^2\,) \\ &= E(\, [1/n \sum_i \lambda_i z_i^2]^2 - [1/n \sum_i \lambda_i]^2\,) \\ &= 1/n^2 \sum_i \sum_j (\, \lambda_i \lambda_j (\, E(z_i^2 z_j^2) - 1\,)\,). \quad \square \end{aligned}$$

Si par exemple on sait que A est tel que ces valeurs propres soient inférieures à 1 (c'est le cas pour les opérateurs de type filtrage), on en déduit immédiatement le majorant suivant pour l'écart type de g(w):

Proposition 2.2.: Si $\max(|\lambda_i|) \leq 1$, alors

$$\sigma(g(w)) \leq \sqrt{2/n}\,.$$

Si A est de plus semi-définie positive alors:

$$\begin{aligned} \sigma(g(w)) &\leq \sqrt{2/n}\ [1/n \mathrm{tr}(A)]^{1/2} \\ &\leq \sqrt{2/n}\,, \end{aligned}$$

la première inégalité étant une égalité si A est une projection.

On peut donc déjà proposer pour le calcul de $1/n \mathrm{tr}(A)$ l'algorithme de Monte-Carlo suivant:

- générer un vecteur w pseudo-aléatoire de loi G(0,I)
- calculer g(w) qui donnera alors une approximation de μ_1.

Les majorations de la proposition 2.2 nous assurent que cette approximation de μ_1 aura une précision absolue d'autant meilleure que la matrice est grande ; en effet rappelons que la précision d'un estimateur X de m, sans biais, est fortement liée à sa variance σ^2 :par exemple par l'inégalité de Tchebychev: $P\{\, |X-m|/\sigma \geq \sqrt{1/(1-\alpha)}\,\} \leq (1-\alpha)$ qui implique que la demi-longueur de l'intervalle de confiance au niveau α est inférieure à $\sigma/\sqrt{1-\alpha}$.

Evidemment suivant l'usage classique des méthodes de Monte-Carlo (où il est rare qu'une seule évaluation suffise), la répétition (M fois) de l'évaluation de g(w) pour plusieurs vecteurs pseudo-alétoires fournirait une erreur $\sqrt{M}$ fois plus petite.

Mais il est assez remarquable qu'ici, avec la seule hypothèse que les valeurs propres de A sont comprises entre 0 et 1, on soit déjà assuré que l'algorithme précédent avec une seule évaluation donne une bonne précision relative si $1/n\ tr(A)$ a une valeur raisonnable (c'est-à-dire pas trop petite). Par exemple si $1/n\ tr(A)$ est voisin de 0.5, l'écart type est majoré par $\sqrt{2/n}\sqrt{0.5}=\sqrt{1/n}$, et donc si n=400, alors $\sigma(g(w))\leq 0.05$, et dans le cas d'un problème de traitement d'images à 200 lignes, 200 colonnes, on a $\sigma(g(w))\leq 0.005$.

Notons que sur les exemples traitées il s'est avéré que la majoration de la proposition 2.2 est plutôt pessimiste pour les faibles valeurs de cette trace . On peut en fait démontrer (cf. [4]) par exemple que si A est la matrice du lissage de n données par une méthode du type spline cubique à pas équidistant, le rapport $[1/n tr(A^2)]^{1/2} / [1/n tr(A)]$ reste compris entre 1 et 5/3 quelque soit le paramètre de lissage.

Plus précisément l'écart type relatif de g(w) est donné par:

$$\sigma(g(w))/E(g(w)) = (\sqrt{2/n})\ \mu_2/\mu_1 = \sqrt{2/n}\ [(1 + d^2/\mu_1^2)]^{1/2}$$

et est donc d'autant plus petit que les λ_i sont peu dispersées.

Notons qu'un estimateur de μ_2^2 est donné par $1/n\ (Aw)^t Aw$ (appliquer la proposition 2.1 avec A^2 au lieu de A puisque $[\mu_2(A)]^2=\mu_1(A^2)$): celui-ci permet d'estimer la précision de g(w) et donc indique le nombre M de tels estimateurs indépendants éventuellement nécessaires pour réduire (par moyenne) ce risque d' erreur.

3. Normalisation par $w^t w$

Si A valait simplement μI, alors l'écart type de g(w) serait $(\sqrt{2/n})|\mu|$. Or dans ce cas trivial la correction de g(w) par le facteur $n/w^t w$ donnerait exactement la valeur de la moyenne arithmétique cherchée.

Cette correction conduit en fait à une meilleure précision sur une matrice quelconque, puisqu'on a la:

Proposition 3.1.: Avec les mêmes notations qu'à la proposition 2.1.; soit w un vecteur de n variables aléatoires indépendantes, de loi gaussienne centrée, de variance 1, (i.e. $w \rightsquigarrow G(0,I)$).
Alors $s(w) = (w^tAw)/(w^tw)$ est un estimateur sans biais de μ_1 d'écart type $\sqrt{2/(n+2)}\, d$, c'est-à-dire:

$$E(s(w)) = 1/n\, tr(A) = \mu_1$$

$$\sigma(s(w)) = \sqrt{2/(n+2)}\ \left[1/n\, tr(A^2) - (1/n\, tr(A))^2 \right]^{1/2} = \sqrt{2/(n+2)}\, d.$$

démonst. : En notant que $s(w) = \sum \lambda_i x_i^2$ où $x = (1/\sqrt{w^t w})Qw$ est un point de la surface sphérique unité S^{n-1}, distribué avec une densité uniforme (puisque pour toute transformation orthogonale U, $Ux = (1/\sqrt{(UQw)^t UQw})UQw$ a même loi que $(1/\sqrt{(Qw)^t Qw})Qw = x$), la démonstration résulte des intégrales (par ex. [5, formule 4.644]):

$$(3.1) \qquad \int x_i^k\, c_n d\omega(x_1,x_2,\ldots,x_n) = \int_0^{\pi} \cos^k t\, \sin^{n-2} t\, c'_n\, dt, \qquad \text{pour tout } i,$$

où $d\omega$ est l'élément de surface sur S^{n-1} et c_n, c'_n normalisent $\int c_n d\omega(x) = 1$, et de la récurrence connue (par ex. [5, formule 2.510]) suivante:

$$(3.2) \qquad \int_0^{\pi} \cos^k t\, \sin^{n-2} t\, dt \;=\; (k-1)/(n+k-2) \int_0^{\pi} \cos^{k-2} t\, \sin^{n-2} t\, dt.$$

En effet si l'on note $v_{i,j} = \int x_i x_j\, c_n d\omega(x_1,x_2,\ldots,x_n)$
et $v_{i,j,k,l} = \int x_i x_j x_k x_l\, c_n d\omega(x_1,x_2,\ldots,x_n)$, on a, d'après (3.1) et (3.2) avec k=4:

$$v_{i,i,i,i} = 3/(n+2)\, v_{i,i}$$

et comme $\int (x_1^2 + \ldots + x_n^2)\, c_n d\omega(x_1,x_2,\ldots,x_n) = 1$ entraîne que pour tout i, $n\, v_{i,i} = 1$, on a ensuite:

$$v_{i,i} = 1/n \qquad , \quad v_{i,i,i,i} = 3/(n+2)n.$$

De plus $\int (x_1^2+\ldots+x_n^2)^2 \, c_n d\omega(x_1,x_2,\ldots,x_n) = 1$ entraîne que pour tout i,j, $i\neq j$, $n\, v_{i,i,i,i} + n(n-1)\, v_{i,i,j,j} = 1$, donc:

$$\text{si } i\neq j, \qquad v_{i,i,j,j} = {}^{1}/_{(n+2)n} .$$

Maintenant on déduit de ces moments que:

$$E(s(w)) = \int \sum_i \lambda_i x_i^2 \, c_n d\omega(x_1,x_2,\ldots,x_n)$$
$$= \sum_i \lambda_i v_{i,i} = {}^{1}/_{n} \sum_i \lambda_i ,$$

et:

$$\sigma^2(s(w)) = E(s(w)^2 - (E(s(w))^2)$$
$$= \int [\sum_i \lambda_i x_i^2]^2 \, c_n d\omega(x_1,x_2,\ldots,x_n) - [{}^{1}/_{n} \sum_i \lambda_i]^2$$
$$= \sum_i \sum_j (\lambda_i\lambda_j (v_{i,i,j,j} - {}^{1}/_{n^2}))$$
$$= \sum_i (\lambda_i^2 ({}^{3}/_{(n+2)n} - {}^{1}/_{n^2})) + \sum_i \sum_{j/j\neq i} (\lambda_i\lambda_j ({}^{1}/_{(n+2)n} - {}^{1}/_{n^2}))$$
$$= {}^{2}/_{(n+2)n^2} [\sum_i (\lambda_i^2 (n-1) - \sum_i \sum_{j/j\neq i} (\lambda_i\lambda_j)]$$
$$= {}^{2}/_{(n+2)} \, d^2,$$

car $d^2 = {}^{1}/_{n} \sum_i (\lambda_i - \mu_1)^2 = {}^{1}/_{n} \sum_i \lambda_i^2 - [{}^{1}/_{n} \sum_i \lambda_i]^2$

$$= {}^{1}/_{n^2} [\sum_i n\lambda_i^2 - \sum_i \lambda_i^2 - \sum_i \sum_{j/j\neq i} (\lambda_i\lambda_j)]. \qquad \square$$

Notons que l'écart type relatif de s(w) est alors $\sqrt{{}^{2}/_{(n+2)}} \, {}^{d}/_{\mu_1}$, c'est-à-dire $\sqrt{{}^{2}/_{(n+2)}}$ fois la dispersion relative des valeurs propres.

En l'absence d'information sur la dispersion d, on a aussi la majoration suivante facile à établir:

Proposition 3.2.: Si A est semi-définie positive et $\max(\lambda_i) \leq 1$, alors

$$\sigma(s(w)) \leq \sqrt{2/(n+2)} \quad [\,1/n\,tr(A)\,[\,1-1/n\,tr(A)]\,]^{1/2}$$

$$\leq 0.5\,\sqrt{2/(n+2)}\,,$$

la première inégalité étant une égalité si A est une projection.

Rappelons que lors du calcul de l'intégrale v d'une fonction d'une variable par une méthode de Monte-Carlo classique, le majorant de l'écart type du résultat d'un seul tirage aléatoire est $\sqrt{v(1-v)}$.
Le majorant de $\sigma(s(w))$ se compare donc très avantageusement au cas classique.

La méthode de Monte-Carlo proposée ici consiste à générer un vecteur w pseudo-aléatoire de loi G(0,I) puis à calculer s(w) qui donnera une approximation de μ_1 d'autant meilleure que la matrice est grande et que la dispersion de ses valeurs propres est faible. On pourrait aussi directement générer un point de la surface de la sphère de $\mathbb{R}^n$ tiré suivant une loi uniforme.

Au sujet de la répétition de l'évaluation de g(w) pour plusieurs vecteurs pseudo-alétoires, notons qu'un estimateur de $d^2=\mu_2{}^2-\mu_1{}^2$ est donné par $[(1/n(Aw)^tAw)-(1/n w^tAw)^2]$: celui-ci permet d'estimer la précision de s et donc indique le nombre M de tels estimateurs indépendants éventuellement nécessaires pour réduire ce risque d'erreur. La comparaison de $\sigma(g(w))$ et $\sigma(s(w))$ montre que le gain peut être très important par rapport à la première méthode.

4. Techniques de réduction de variance

Si A est une matrice "voisine" de βI , la première méthode appliquée avec $A-\beta I$ à la place de A, c'est-à dire

$$\begin{aligned} 1/n\ tr(A) &= 1/n\ tr(A-\beta I) + 1/n\ tr(\beta I) \\ &\approx 1/n\ w^t(A-\beta I)w + \beta \\ &\approx 1/n\ w^tAw - \beta/n\ w^tw + \beta, \end{aligned}$$

devrait être utilisée de préférence dès que l'écart type $\mu_2(A-\beta I)$ de cet estimateur est inférieur à $\mu_2(A)$. On remarque que cet écart type serait évidemment minimum pour $\beta=\mu_1$.

La 2ème méthode est, elle, insensible à la translation précédente des valeurs propres. Mais s'il existe une matrice B de trace connue, dont les valeurs propres sont voisines de A, et si l'on sait calculer le produit By pour un y donné, on a intérêt à prendre comme estimateur:

$$1/n\ tr(A) \approx 1/n\ tr(B) + (w^tAw - w^tBw)/w^tw,$$

puisque l'écart type de cet estimateur est la dispersion d(A-B) des valeurs propres de A-B.

5. Un exemple d'application

Nous considérons ici la résolution d'un système linéaire associée à la minimisation d'une forme quadratique donnée. Le problème est celui de la cartographie de données bruitées: supposons que l'on connaisse des valeurs approchées $y_{i,j} \approx g(i\cdot h, j\cdot h)$ d'une fonction g inconnue supposée lisse, aux noeuds (i,j), $i=1,\cdots,n_1$, $j=1,\cdots,n_2$ (où h est le pas >0) d'une grille couvrant le domaine $[0,n_1\cdot h]x[o,n_2\cdot h]$, et que, par exemple, les erreurs des $y_{i,j}$ soient centrées et de même variance.

Comme "reconstruction" en ces points de la fonction g, nous choisissons le tableau $f_{\tau,i,j}$ $i=1,..,n_1, j=1,n_2$, qui, pour un paramètre donné τ, minimise

$$\sum_{i=1}^{n}\sum_{j=1}^{n}(f_{i,j}-y_{i,j})^2 + \tau\sum_{i=2}^{n-1}\sum_{j=2}^{n-1}\Big[\ [(f_{i+1,j}-2f_{i,j}+f_{i-1,j})/h^2]^2 + 2[((f_{i+1,j+1}-f_{i+1,j-1})-(f_{i-1,j+1}-f_{i-1,j-1}))/4h^2]^2 + [(f_{i,j+1}-2f_{i,j}+f_{i,j-1})/h^2]^2\ \Big],$$

sur l'ensemble des tableaux f. Cette approche est une extension à deux dimensions de la méthode spline de Whittaker. En représentant chaque tableau $n_1 \times n_2$ par un vecteur à $n_1\cdot n_2$ composantes, (avec l'ordre lexicographique), il est facile de voir qu'il existe une matrice Ω d'ordre $n=n_1\cdot n_2$, symétrique, pentadiagonale par blocs, où les blocs sont d'ordre n_1, telle que f_τ est solution du système linéaire $(I + \tau\ \Omega)\ f_\tau = y$.

La résolution directe est à écarter: la matrice est certes bande, et creuse (précisément 13 diagonales), mais de largeur $4\,n_1 + 1$, et la décomposition de Cholesky demanderait donc $O(n_1{}^3 n_2)$ opérations et,

surtout, de stocker une bande de $(2n_1+1)\cdot n_1\cdot n_2$ éléments. Par contre les itérations du gradient conjugué sont peu coûteuses (chaque itération consiste principalement en un produit par cette matrice creuse , qui demande au plus 13 $n_1\cdot n_2$ opérations et qui n'a pas besoin que cette matrice soit stockée) et convergent toujours en une dizaine d'itérations pour les lissages courants (si τ devient grand, le problème devient mal conditionné).

On présente les expériences réalisées pour l'estimation de la trace de $I-(I+\tau\,\Omega)^{-1}$ avec $n_1=30$, $n_2=46$, et $\tau/h^4=.1$ qui correspond à un lissage courant sur cette géométrie. On a répété 5 fois la génération d'un bruit blanc w suivie de la résolution du système linéaire avec w comme second membre et de l'évaluation des 2 estimateurs proposés. Ces résultats ont été obtenus en simple précison sur un VAX 730, sous VMS, avec son générateur de loi uniforme, l'échantillon Gaussien en étant déduit par la méthode directe (avec l'approximation rationnelle classique [1, formule 26.2.23] pour l'inverse de la fonction de répartition).
Ces estimateurs sont à comparer avec la trace 'exacte' valant ici: $1/n\,tr(A)=0.439239$, qui été obtenue par la résolution des $n_1\cdot n_2$ systèmes avec les vecteurs canoniques pour second membre (en fait ici la résolution d'un quart de ces systèmes a été suffisante en raison de la symétrie du problème par rapport au centre de la grille). On a chaque fois utilisé la même méthode de gradient conjugué (le test d'arrêt étant que la dernière correction soit inférieure à 10^{-4} fois le second membre, en moyenne quadratique).

calcul de $1/n\,tr(A)$ où $A = I-(I+\tau\,\Omega)^{-1}$ avec $n_1=30$, $n_2=46$, et $\tau/h^4=1$

$\hat{\mu}_2=1/n\,(Aw)^tAw$, et $\hat{\sigma}^2=(1/n(Aw)^tAw)-(1/n w^tAw)^2$

g(w)	$[2/n\,\hat{\mu}_2]^{1/2}$	s(w)	$[2/n+2\;\hat{\sigma}^2]^{1/2}$
0.4486082	0.0191136	0.4344937	0.0086334
0.4492577	0.0191025	0.4338101	0.0085075
0.4366112	0.0189841	0.4446426	0.0084452
0.4620579	0.0194424	0.4399014	0.0082821
0.4435612	0.0192712	0.4428376	0.0092865

Ces valeurs montrent la bonne précision des estimateurs de la trace. Avec l'estimateur normalisé s(w), on obtient, à 1% près, la trace en un seul tirage. Ajoutons que la valeur de la trace dite exacte obtenue plus haut ne peut avoir, elle-même, qu'au plus 3 ou 4 chiffres significatifs en raison de l'accumulation ($n_1\cdot n_2$ fois) des erreurs commises à chaque

résolution.
On note dans les 2ème et 4ème colonnes du tableau que les écarts types sont eux aussi bien estimés en une seule évaluation.
Ces résultats indiquent que la majoration de la proposition 2.1 (applicable puisque, Ω étant semi-définie positive, les valeurs propres de $I+\tau\Omega$ sont supérieures à 1) est un peu pessimiste ici. En fait on peut établir pour ce problème une majoration de l'écart relatif, valable quelque soit $\tau>0$, qui est bien meilleure pour les faibles valeurs de $tr(I-(I+\tau\Omega)^{-1})$ [4].

6. Conclusions

Cette méthode de calcul d'une trace, qui à la connaissance de l'auteur n'avait pas encore été décrite, est donc une méthode pouvant être d'une très bonne précision pour de grandes matrices.

Puisqu'elle ne nécessite qu'un seul produit Ay aux lieu de n dans la méthode classique, le gain en coût est important pour les grandes matrices. Certes dans le cas où A est l'inverse d'une matrice Z donnée et où le calcul de Ay (une résolution) coûte $O(n^3)$ opérations, le coût des n résolutions est réduit de n^4 à environ $4n^3$ si on factorise Z (par exemple par la méthode de Cholesky), et le gain n'est plus qu'un facteur 4, mais la factorisation, qui doit être stable (pour cette raison on préfère souvent la décomposition en valeurs propres), est très coûteuse en mémoire pour de grandes matrices creuses.

Notons que si A est donnée seulement sous la forme du produit $A_pA_{p-1}...A_1$ de p (≥ 3) matrices (comme cela apparaît dans les méthodes itératives stationnaires), cette méthode peut apporter aussi une solution remarquablement économique dès que l'on sait calculer le produit A_ky pour $1\leq k\leq p$ et pour tout y donné: par exemple si les A_k sont des matrices pleines données, chaque élément i,j de la matrice produit est la somme des n^{p-1} termes $A_{p,i,i_{p-1}}A_{p-1,i_{p-1},i_{p-2}}...A_{1,i_1,j}$, $1\leq i_k\leq n$, $1\leq k\leq p-1$, et donc la trace exacte peut se calculer en $(p-1)n^p$ opérations (la factorisation du produit peut évidemment réduire ce coût à $(p-2).n^3+n^2$ au prix de la mémorisation successive des p-2 matrices A_2A_1, $A_3A_2A_1$,..., $A_{p-1}...A_2A_1$), alors que l'approximation proposée ici se ramène aux évaluations successives de $w_1=A_1w$, $w_2=A_2w_1$,..., $w_p=A_pw_{p-1}$ plus le produit scalaire w^tw_p, c'est-à-dire seulement $p.n^2+n$ opérations. Si l'une ou plusieurs des matrices A_k ne sont données que sous la forme de l'inverse d'une matrice donnée, le gain peut évidemment être encore plus important.

Parmi les applications possibles, citons le calcul de statistiques en

traitement de données, permettant par exemple la mise en oeuvre de la méthode de validation croisée pour le choix de paramètres de lissage. Cette application est étudiée dans [4].

La méthode proposée permet aussi d'approcher un majorant du conditionnement $\kappa(Z)$ d'une grande matrice symétrique Z donnée, si on prend la norme Euclidienne pour normer les vecteurs. En effet $\kappa(Z)$, alors défini par $\kappa(Z)=\|Z\|\,\|Z^{-1}\|$ où $\|A\|=\sup_x(\|Ax\|/\|x\|)=\max_i \lambda_i(A)$, a un majorant déduit de $\lambda_i(A)\leq n\mu_1(A)=tr(A)$, qui est le produit $tr(Z)\,tr(Z^{-1})$ et qui peut donc estimé au coût d'une résolution du système $Zx=w$ (notons que pour n grand, l'inégalité $\lambda_i-\mu_1\leq\sqrt{n}\,d$ donnera un bien meilleur majorant de $\kappa(Z)$ si les valeurs propres de Z^{-1} sont peu dispersées).

Remarquons enfin que si A n'est pas symétrique, rien n'interdit d'utiliser les algorithmes présentés ici, et en fait, comme $w^t Aw = w^t A^t w = {}^1/_2 w^t(A+A^t)w$, les résultats précédents indiquant l'erreur alors commise, restent valables avec ${}^1/_2(A+A^t)$ à la place de A.

Références

[1] **Abramowitz M., Stegun I.A.** Handbook of mathematical functions. Dover Publications, Inc., NY. (1972)

[2] **Erisman A. M., Tinney W. F.** : On computing certain elements of the inverse of a sparse matrix. Commun. ACM 18,177-179 (1975)

[3] **Girard D.** Optimal regularized reconstruction in computerized tomography, preprint (1985), à paraître (1987) dans SIAM J. Sci. Statist. Comp.

[4] **Girard D.** A fast and efficient procedure for generalized cross-validation with large data sets. Rapport de recherche TIM3 (Mai 1987).

[5] **Gradshteyn I. S., Ryzhik I. M.** Table of integrals, series, and products. Alan Jeffrey, eds., Academic Press (1980)

[6] **Utreras F.** : Sur le choix de paramètre d'ajustement dans le lissage par fonctions spline. Numer. Math. 34, 15-28 (1980)

[7] **Wahba G.** :Constrained regularization for ill-posed linear operator equations, with applications in meteorology and medicine. In Statistical Decision Theory and related topics, III, Vol. 2, S. S. Gupta and J. O. Bergen, eds., Academic Press (1982)