ANALYSE NUMÉRIQUE ET PROBABILITÉS.

NOMBRES DE CONDITION ET MATRICES ALÉATOIRES.

Mario Wschebor

Centro de Matemática. Facultad de Ciencias

Universidad de la República.

Montevideo. URUGUAY.

wschebor@cmat.edu.uy

---

1 NOMBRE DE CONDITION D'UNE MATRICE CARRÉE

$$Ax = b \qquad (1)$$

Système linéaire $n \times n$

A matrice réelle

$b \in \mathbb{R}^n$

$\|x\|$ = norme euclidienne dans $\mathbb{R}^n$

$$\|A\| = \sup_{\|x\|=1} \|Ax\|$$

A non-singulière $\boxed{\mathcal{K}(A) = \|A\|\|A^{-1}\|}$

A singulière $\mathcal{K}(A) = +\infty$

[Turing (1948) ; Von Neumann-Goldstine (1947)]

$\mathcal{K}(A)$ est une mesure de la propagation de l'erreur relatif entre les données de (1) et la solution.

$$\frac{\|\Delta x\|}{\|x\|} \leq \mathcal{K}(A)\left[\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|}\right] + o\left(\frac{\|\Delta A\|}{\|A\|} + \frac{\|\Delta b\|}{\|b\|}\right)$$

$\mathcal{K}(A)$ est la plus petite constante avec cette propriété au premier ordre.

$\mathcal{K}(A)$ figure de façon naturelle dans l'étude de certaines algorithmes.

EXAMPLE : A symmétique, définie positive
Méthode de Cholesky pour résoudre (1)
Alors :

$$\frac{\|\Delta x\|}{\|x\|} \leqq 3\, n^3\, u\, \mathcal{K}(A) \qquad (2)$$

$\underline{\underline{u}}$ est l'arrondi de la machine, p.e. $u = 2^{-\ell}$

---

(2) montre une situation générale : l'erreur relative de la solution va dépendre de :

* la taille de l'entrée
* la précision de la machine
* le conditionnement du problème à résoudre

---

* RANDOMISATION DE **A**.

* ÉTUDE DE LA LOI DE $\varkappa(A)$
  [OU DE $\log \varkappa(A)$]

PROPRIÉTÉS SIMPLES DE $\varkappa(A)$

* $$\varkappa(A) = \frac{\|A\|}{d_F(A, \Sigma)}$$

  Σ = ensemble des matrices singulières

  d_F = distance de Frobenius : $d_F(A,B) = \left[\sum_{i,j} (a_{ij} - b_{ij})^2\right]^{1/2}$

* $\lambda_1, \dots, \lambda_n$ valeurs propres de $A^T A$

  $$0 \leq \lambda_1 \leq \dots \leq \lambda_n$$

  $$f(x) = \|Ax\|^2 = x^T A^T A x$$

  $$\lambda_n = \|A\|^2 = \max_{\|x\|=1} f(x)$$

  $$\lambda_1 = \frac{1}{\|A^{-1}\|^2} = \min_{\|x\|=1} f(x) \qquad [\text{si } \lambda_1 > 0]$$

  $$\Rightarrow \qquad \varkappa(A) = \left(\frac{\lambda_n}{\lambda_1}\right)^{1/2}$$

* $\varkappa(A) \geq 1$

  $\varkappa(cA) = \varkappa(A) \quad \forall\, c \in \mathbb{R},\ c \neq 0.$

LORSQUE **A** EST ALÉATOIRE, ON A UN PROBLÈME SUR LE SPECTRE DES MATRICES ALÉATOIRES

THEORÈME (Edelman, 1988, SIAM J. Matrix Anal.)

$$A = ((a_{ij}))_{i,j=1,\dots,n}$$

a_{ij} i.i.d. $N(0,1)$

$$\Rightarrow E\{\log \mathcal{K}(A)\} = \log n + C_0 + \varepsilon_n$$

- C_0 constante connue, $C_0 \cong 1.537$
- $\varepsilon_n \longrightarrow 0$

Idée de la démonstration:

$$\log \mathcal{K}(A) = \frac{1}{2}\log \lambda_n - \frac{1}{2}\log \lambda_1$$

DANS CE CAS, on a une formule pour la densité jointe de $\lambda_1, \dots, \lambda_n$ depuis Fisher et alt (1939). On peut en déduire le comportement des marginales de λ_1 et λ_n lorsque $n \to \infty$.

Démonstration analytique (fonctions spéciales, etc.)

THÉORÈME (Cas non-centré, J. Cuesta[illegible], 200[illegible], IV Ascona Seminar, Birkhäuser).

a_{ij} v.a. indép. ; $a_{ij} \sim \quad N(m_{ij}, \sigma^2)$

$M = ((m_{ij}))$

$$\Rightarrow E\{\log \mathcal{K}(A)\} \leq \log n + \log\left(\frac{\|M\|}{\sigma\sqrt{n}} + 4\right) + C_1$$

C_1 constante connue.

ON PEUT APPLIQUER LA THÉORIE SUR LE SPECTRE DES MATRICES ALÉATOIRES (INÉGALITÉS DE CONCENTRATION, ETC) À L'ÉTUDE DE $\mathcal{K}(A)$.

RÉFÉRENCES :

* SZAREK, 1991, J. of Complexity
* DAVIDSON-SZAREK, Handbook on the Geometry of Banach Spaces, Ch. 8, 2001
* LEDOUX, 2002.

On obtient, par exemple :

$$P\left(\frac{1}{n}\mathcal{K}(A) > x\right) \leq (\text{const})\, \frac{1}{x}\left(\frac{\log x}{n}\right)^{\alpha}$$

α une constante positive.

(ceci dans le cas a_{ij} i.i.d. $N(0,1)$)

a_{ij} i.i.d. $N(0,1)$, $n \geq 3$.

$$\Longrightarrow \quad \frac{c}{x} < P\left(\frac{1}{n}\chi(A) > x\right) < \frac{C}{x} \qquad \forall x > 1$$

c, C constantes positives.

Idée de la démonstration

À différence des résultats précédents, la démonstration dépend d'une estimation fine de la densité jointe

$$g(a,b), \quad a > b$$

de λ_n, λ_1.

* $B = A^T A = ((b_{ij}))_{i,j=1,\dots,n}$

* $X : S^{n-1} \longrightarrow \mathbb{R}$

  $X(x) = x^T B x$

* $\Pi_s : \mathbb{R}^n \longrightarrow \mathbb{R}^n$ proj. orth. sur $s^\perp$ $(s \neq 0)$

* $M \succ 0$ (resp. $M \prec 0$) veut dire que M est symétrique déf. posit. (resp. déf. nég.).

$a > b$

$$\{\lambda_m \in (a, a+da),\ \lambda_1 \in (b, b+db)\} =$$

$$= \Big\{ \exists\, s,t \in S^{m-1},\ \langle s,t\rangle = 0,\ X(s) \in (a, a+da),\ X(t) \in (b, b+db),$$
$$\Pi_s(Bs) = 0,\ \Pi_t(Bt) = 0,\ X''(s) < 0,\ X''(t) > 0 \Big\} \circledast$$

$N_{a,b,da,db}$ = nombre des couples (s,t) vérifiant $\circledast$

$= 0$ ou 4

$$\Rightarrow \quad g(a,b)\,da\,db = \frac{1}{4} E\{N_{a,b,da,db}\}$$

LE POINT PRINCIPAL VIENT DU FAIT que l'on peut calculer l'éspérance par une formule de type Rice.

variété C^∞ sans bord, compacte, dim $(V) = 2n-3$

* $\tau = (s,t) \in V$

  $\sigma_V(d\tau)$ mesure géométrique sur V

  $\sigma_V(V) = \sqrt{2}\,\sigma_{n-1}\,\sigma_{n-2}$ $\left[\sigma_{n-1} = \text{aire de } S^{n-1} = \dfrac{2\pi^{n/2}}{\Gamma(\frac{n}{2})}\right]$

* $Y : V \longrightarrow \mathbb{R}^{2n}$ champs vectoriel aléatoire

$$Y(s,t) = \begin{pmatrix} \pi_s(Bs) \\ \pi_t(Bt) \end{pmatrix}$$

* <u>FORMULE DE RICE</u>

$$E\{N_{a,b,da,db}\} = \int_a^{a+da} dx \int_b^{b+db} dy \int_V E\{\Delta(s,t)\, 1_{\{X''(s)<0,\, X''(t)>0\}}$$

$$/ X(s)=x,\ X(t)=y,\ Y(s,t)=0\}\, p_{X(s),X(t),Y(s,t)}(x,y,0)\ \sigma_V(d(s,t))$$

où

* $\Delta(\tau) = \left[\det\left[(Y'(\tau))^T\, Y'(\tau)\right]\right]^{1/2}$

* $p_{X(s),X(t),Y(s,t)}$ densité du triplet dans $\mathbb{R}^1 \times \mathbb{R}^1 \times W_{s,t}$

  $W_{s,t} = \{(t,-s)\}^\perp \cap \{s^\perp \times t^\perp\}$ dim $2n-3$

Invariance par rotations de $\mathbb{R}^n$ de la loi de A

$\Rightarrow$ l'intégrand ne dépend pas de (s,t)

$\Rightarrow$

$$g(a,b) = \frac{\sqrt{2}}{4}\,\sigma_{n-1}\,\sigma_{n-2}\, E\Big\{ \Delta(e_1,e_2)\, \mathbb{1}_{\{X''(e_1) < 0,\, X''(e_2) > 0\}} \Big/ X(e_1)=a,\, X(e_2)=b,\, Y(e_1,e_2)=0 \Big\}.$$

$$\cdot\; p_{X(e_1),\, X(e_2);\, Y(e_1,e_2)}(a,b,0).$$

Le reste c'est des calculs élémentaires.

— • —

QUELQUES PROBLÈMES CONCERNANT $\mathcal{K}(A)$

* Changement de loi.

* Matrices structurées.

  Example (Viswanath – Trefethen, 1998)

$$\left.\begin{array}{l} a_{ij} = 0 \quad \text{si} \quad i<j \\ a_{ij} \quad \text{i.i.d} \quad N(0,1) \quad \text{si} \quad i \geq j \end{array}\right\} \Rightarrow \text{p.s.} \quad (\mathcal{K}(A))^{1/n} \longrightarrow 2$$

[régime totalement différent à celui des matrices pleines].

* Matrices rectangulaires

— • —

Est-ce-que

$$Ax < 0 \qquad (3)$$

a des solutions ?

Ici

$$A = ((a_{ij}))_{i=1,\dots,n;\, j=1,\dots,m} \text{ réelle.}$$

COMMENTAIRES

* Dualité
* Algorithmes de point intérieur
* Les cas intéressants sont $n > m$ [et même $n \gg m$]

---

* $A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} \qquad a_k \in \mathbb{R}^m \quad (k=1,\dots,n)$

$$\mathcal{F} = \mathcal{F}_{n,m} = \{ A \text{ t.q. (3) admet une solution} \}$$

$$\Sigma = \partial \mathcal{F}$$

* Renegar (SIAM J. Opt., 1994-95)

$$C_R(A) = \frac{\|A\|}{d_F(A, \Sigma)}$$

$Ax < 0$

$$\left[\|A\| = \sup_{\|x\|_m = 1} \|Ax\|_n \right]$$

* Cheung-Cucker (Math. Program., 2001)

$$Z : \mathbb{R}^m \to \mathbb{R} \quad : \quad Z(x) = \max_{1 \leq k \leq n} \frac{\langle a_k, x \rangle}{\|a_k\|}$$

$\|a_k\|$

$$\underline{Z} = \min_{x \in S^{m-1}} Z(x)$$

$$\boxed{\mathscr{C}(A) = \frac{1}{|\underline{Z}|}}$$

* $\Sigma' = \{ A : \underline{Z} = 0 \}$ $\left[\text{i.e. } \Sigma' = \{ A : \mathscr{C}(A) = +\infty \} \right]$

RANDOMISATION DE **A** : $\frac{a_k}{\|a_k\|}$ i.i.d. uniformes sur S^{m-1}

où a_k i.i.d. $N(0, I_m)$ dans $\mathbb{R}^m$.

$(k = 1, \ldots, n)$.

* Qu'est-ce-qu'on peut dire sur la distribution de $\mathcal{C}(A)$?

i.e. $P(|\underline{\underline{Z}}| < a)$?

Ou bien sur les moments de $\log \mathcal{C}(A)$?

* DIFFICULTÉ MAJEURE :

La géometrie de Σ' dans $M_{n \times m}$.

* RAPPORTS AVEC LA PROBABILITÉ GÉOMÉTRIQUE

$-1 < a < 1$

$P(\underline{\underline{Z}} > a)$ n'est que la probabilité de recouvrir S^{m-1} avec $\underline{\underline{n}}$ calottes de rayon angulaire Arcos a et centres uniformément distribués sur S^{m-1}.

Résultats très partiels :
* P. Hall. "Int. Coverage Processes", 1988
* S. Janson. Acta Math., 1986

se limiter à considérer le cas:

$$\frac{m\ (1+\log n)}{n} \leqq 1$$

<u>THÉORÈME</u> (F. Cucker - MW, Numerische Math., 2003)

* $E\{\log \mathscr{C}(A)\} \leqq \max\{\log m, \log\log n\} + C = L_{n,m}$

* $E\{[\log \mathscr{C}(A)]^{\nu}\} \leqq 2\nu! \sum_{k=0}^{\nu} \frac{1}{k!} L_{n,m}^{k} \qquad (\nu = 2,3,\dots)$

* Pour tout λ, $0<\lambda<1$ on a

$$P\left(\log \mathscr{C}(A) > L_{n,m}\, y\right) \leqq \left(1 + 2e\frac{\lambda}{\lambda-1}\right) e^{-\lambda y} \qquad (y>0)$$

Encore ici, on peut essayer d'appliquer des techniques de champs vectoriels aléatoires.

Pour $a>0$:

$$P(-a < \underline{\underline{Z}} < a) = P\left(m^{\tilde{\underline{\underline{Z}}}}(a) \geqq 1\right) \underset{\uparrow}{\leqq} E\left\{m^{\tilde{\underline{\underline{Z}}}}(a)\right\}$$

où

$$\tilde{\underline{\underline{Z}}} = Z / S^{m-1}$$

$m^{\tilde{\underline{\underline{Z}}}}(a) =$ nombre des minima locaux de $\tilde{\underline{\underline{Z}}}$ dont la valeur est comprise entre $-a$ et a.

L'inégalité $\uparrow$ est fine si $\underline{\underline{a}}$ est petit.

LE PROBLÈME POUR CALCULER $E\{m^{\tilde{\underline{\underline{Z}}}}(a)\}$ est que $\tilde{\underline{\underline{Z}}}$ est très irrégulière. Il faut d'abord la lisser et puis délisser.

ÉCHANTILLON DE QUESTIONS (PRESQUE TOUT EST À FAIRE)

* Changement de loi.
* Problèmes non-linéaires.
