Université de Rennes 1
Préparation à l'agrégation
de mathématiques
Auteur du document :

M. Coste

Comptage de racines et signature de formes quadratiques

Michel Coste

7 octobre 2003

La première section de ce texte présente, avec des références des développements sur les formes quadratiques ou sur les racines de polynômes. La deuxième section fait manipuler des formes quadratiques sur l'algèbre $\mathbb{R}[X]/P$. Je ne connais malheureusement pas de référence pour cette partie dans la bibliothèque de l'agrégation. La troisième section présente quelques méthodes de calcul de signature et un algorithme implémenté en Maple.

1 Deux méthodes

On peut trouver deux références dans la bibliothèque de l'agrégation qui expliquent comment calculer le nombre de racines réelles d'un polynôme au moyen de la signature d'une forme quadratique. Nous présentons brièvement ce qui est dit dans ces références. Dans tout ce qui suit, on se donne un polynôme $P = X^n + a_1X^{n-1} + \cdots + a_n$ à coefficients réels. On note $\alpha_1, \ldots, \alpha_n$ les racines de P (dans $\mathbb{C}$, comptées avec leurs multiplicités).

1.1 L'exposé de Gantmacher

La référence est Gantmacher, Théorie des matrices tome 2 (Dunod), p. 199. Soient $s_0, s_1, s_2 \ldots$ les sommes des puissances des racines de P. On a $s_i = \sum_{k=1}^{n} \alpha_k^i$, et en particulier $s_0 = n$. Soit S la forme quadratique réelle de dimension n définie par

$$S(u) = \sum_{i,j=0}^{n-1} s_{i+j}u_iu_j \ , \text{ où } u = (u_0, \ldots, u_{n-1}) \ .$$

La matrice de cette forme quadratique est

$$H = \begin{pmatrix} s_0 & s_1 & \ldots & s_{n-1} \\ s_1 & s_2 & \ldots & s_n \\ \vdots & \vdots & \vdots & \vdots \\ s_{n-1} & s_n & \ldots & s_{2n-1} \end{pmatrix} \ .$$

Théorème 1 *Soit (s,t) la signature de la forme quadratique S. Le nombre de racines réelles distinctes de P est $s-t$. Le nombre de racines réelles ou complexes distinctes est égal au rang $s+t$.*

La démonstration commence par la remarque que la forme quadratique S se décompose en somme de carrés de formes linéaires sur $\mathbb{C}$ de la manière suivante :

$$S(u) = \sum_{k=1}^{n} (u_0 + \alpha_k u_1 + \alpha_k^2 u_2 + \cdots + \alpha_k^{n-1} u_{n-1})^2 \ .$$

Supposons que P a q racines distinctes et notons $\ell_1, \ldots, \ell_q$ les formes linéaires correspondant à ces différentes racines. On a donc $S = \sum_{k=1}^{q} m_k \ell_k^2$, où les m_k sont les multiplicités des racines. Ces formes linéaires $\ell_1, \ldots, \ell_q$ sont linéairement indépendantes (le déterminant de Vandermonde formé par leurs q premiers coefficients est non nul). Ceci montre déjà que le rang de la forme quadratique est q. Si la forme ℓ_k correspond à une racine complexe non réelle (et $\overline{\ell_k}$ à sa conjuguée), on a, en notant v et w les formes linéaires à coefficients réels telles que $\ell_k = v + iw$ et $\overline{\ell_k} = v - iw$:

$$m_k(\ell_k^2 + \overline{\ell_k}^2) = 2m_k v^2 - 2m_k w^2 \ ,$$

ce qui montre qu'une racine complexe et sa conjuguée contribuent pour $(1,1)$ à la signature. Bien sûr, une racine réelle contribue pour $(1,0)$ à la signature. Le théorème est donc montré.

Une remarque sur le calcul des sommes de puissances des racines. Elles s'obtiennent par récurrence en fonction des coefficients de P, en utilisant les classiques formules de Newton (par exemple : Gourdon, Algèbre, Ellipses, p. 81). Ces formules s'écrivent, pour tout entier strictement positif p :

$$p a_p + \sum_{i=1}^{p} s_i a_{p-i} = 0 \ , \text{ où } a_0 = 1 \text{ et } a_j = 0 \text{ pour } j > n \ . \qquad (N_p)$$

Un moyen commode d'obtenir ces formules consiste à introduire le polynôme réciproque de P

$$Q = X^n P(1/X) = 1 + a_1 X + \cdots + a_n X^n = \prod_{k=1}^{n} (1 - \alpha_k X) \ ,$$

à écrire l'égalité dans l'anneau de séries formelles en X

$$-Q' = Q \left(\sum_{k=1}^{n} \frac{\alpha_k}{1 - \alpha_k X} \right) = Q \left(\sum_{k=1}^{n} \sum_{i=0}^{\infty} \alpha_k^{i+1} X^i \right) = Q \sum_{i=0}^{\infty} s_{i+1} X^i \ ,$$

puis à identifier les coefficients de X^{p-1} des deux côtés pour obtenir (N_p).

1.2 L'exposé de Dieudonné

La référence ici est Dieudonné, Calcul infinitésimal, p. 64. On introduit le polynôme symétrique en X et Y

$$L(P, X, Y) = \frac{P(X)P'(Y) - P(Y)P'(X)}{X - Y} = \sum_{i,j=0}^{n-1} a_{i,j} X^i Y^j \ ,$$

avec les coefficients duquel on fabrique la forme quadratique

$$Q(P,u) = \sum_{i,j=0}^{n-1} a_{i,j} u_i u_j \ , \text{ où } u = (u_0, \ldots, u_{n-1}) \ .$$

La matrice de Q est quelquefois appelée *matrice de Bezout* du polynôme P.

Théorème 2 *Soit (s,t) la signature de la forme quadratique Q. Le nombre de racines réelles distinctes de P est $s-t$. Le nombre de racines réelles ou complexes distinctes est égal au rang $s+t$.*

L'idée de la démonstration donnée par Dieudonné est la suivante. Soit $P_1 = X^p + b_1 X^{p-1} + \cdots + b_p$ et $P_2 = X^q + c_1 X^{q-1} + \cdots + c_q$. Alors

$$L(P_1 P_2, X, Y) = P_2(X) P_2(Y) L(P_1, X, Y) + P_1(X) P_1(Y) L(P_2, X, Y) \ ,$$

et on en déduit que

$$Q(P_1 P_2, u) = Q(P_1, v) + Q(P_2, w)$$

où

$$\begin{cases} v = (v_0, \ldots, v_{p-1}) & \text{avec } v_i = u_{i+p} + b_1 u_{i+p-1} + \cdots + b_p u_i \\ w = (w_0, \ldots, w_{q-1}) & \text{avec } w_i = u_{i+q} + c_1 u_{i+q-1} + \cdots + c_q u_i \end{cases} .$$

Le déterminant des formes linéaires $v_0, \ldots, v_{p-1}, w_0, \ldots, w_{q-1}$ est au signe près le résultant de P_1 et P_2. Donc, si ces polynômes sont premiers entre eux, la signature de $Q(P_1, P_2, u)$ est la somme de celle de $Q(P_1, v)$ et de celle de $Q(P_2, w)$. On se ramène ainsi à étudier le cas des polynômes $P = (X-\alpha)^m$ où α est réel et $P = ((X-a)^2 + b^2)^m$, où a et b sont réels, $b \neq 0$. Dans le premier cas on obtient

$$L(P, X, Y) = m(X-\alpha)^{m-1}(Y-\alpha)^{m-1} \ ,$$

et la signature de Q est $(1,0)$. Dans le deuxième cas on obtient

$$L(P, X, Y) = 2m((X-a)^2 + b^2)^{m-1}((Y-a)^2 + b^2)^{m-1}((X-a)(Y-a) - b^2) \ ,$$

et la signature de Q est $(1,1)$. Ceci permet de conclure.

En fait le résultat montré par Dieudonné est différent : dans la définition de $L(P,X,Y)$: il pose, pour $a \in \mathbb{R}$,

$$L(P, X, Y) = \frac{P(X)P'(Y)(Y-a) - P(Y)P'(X)(X-a)}{X-Y}$$

et il obtient que $s-t$ est le nombre de racines réelles distinctes de P strictement plus grandes que a moins le nombre de racines réelles distinctes de P strictement plus petites que a. Mais la démonstration suit exactement le chemin indiqué ci-dessus.

2 Les deux ne font qu'un

On ne voit pas a priori le rapport entre les deux méthodes exposées ci-dessus. Nous allons montrer qu'il s'agit en fait dans les deux cas de la même forme quadratique, lue dans deux bases différentes.

2.1 L'algèbre $\mathbb{R}[X]/P$

Les éléments de l'algèbre quotient $\mathbb{R}[X]/P$ peuvent être représentés de manière unique par un polynôme de degré strictement plus petit que n (la classe de Q est représentée par le reste de la division euclidienne de Q par P). Nous identifierons donc les éléments de $\mathbb{R}[X]/P$ avec les polynômes de degré $< n$.

L'algèbre $\mathbb{R}[X]/P$ admet la base $(1, X, \ldots, X^{n-1})$ en tant qu'espace vectoriel sur $\mathbb{R}$. La matrice dans cette base de l'endomorphisme de multiplication par X est la matrice compagnon du polynôme P :

$$C_P = \begin{pmatrix} 0 & 0 & \ldots & \ldots & 0 & -a_n \\ 1 & 0 & \ddots & & \vdots & -a_{n-1} \\ 0 & 1 & \ddots & \ddots & \vdots & \vdots \\ \vdots & \ddots & \ddots & \ddots & 0 & \vdots \\ \vdots & & \ddots & \ddots & 0 & -a_2 \\ 0 & \ldots & \ldots & 0 & 1 & -a_1 \end{pmatrix} .$$

Le polynôme caractéristique de cette matrice est P, et donc les valeurs propres sur $\mathbb{C}$ de la multiplication par X sont les racines $\alpha_1, \ldots, \alpha_n$ de P (comptées avec multiplicité). Si A est un élément de $\mathbb{R}[X]/P$, les valeurs propres de l'endomorphisme de multiplication par A sont donc $A(\alpha_1), \ldots, A(\alpha_n)$. La *trace* de A, que l'on définit comme la trace de cet endomorphisme de multiplication par A, est donc $\mathrm{trace}(A) = \sum_{k=1}^{n} A(\alpha_k)$. L'application $A \mapsto \mathrm{trace}(A)$ est une forme linéaire sur $\mathbb{R}[X]/P$.

L'application $A \mapsto \mathrm{trace}(A^2)$ est une forme quadratique T sur $\mathbb{R}[X]/P$. Pour i et j entre 0 et $n-1$, on a $\mathrm{trace}(X^i X^j) = \sum_{k=1}^{n} \alpha_k^{i+j} = s_{i+j}$. Ceci montre :

Proposition 1 *la matrice H de la section 1.1 est la matrice de la forme T dans la base $(1, X, \ldots, X^{n-1})$.*

On peut introduire une autre forme linéaire $\ell : \mathbb{R}[X]/P \to \mathbb{R}$ définie par $\ell(X^{n-1}) = 1$ et $\ell(X) = 0$ pour $0 \leq i < n-1$. La forme trace s'exprime au moyen de la forme ℓ par

Lemme 1 *Pour tout élément A de $\mathbb{R}[X]/P$, on a* $\mathrm{trace}(A) = \ell(P'A)$.

Pour montrer cette relation, on étend les scalaires à $\mathbb{C}$ et on écrit

$$P'A = A \sum_{k=1}^{n} \frac{P}{X - \alpha_k} \equiv \sum_{k=1}^{n} A(\alpha_k) \frac{P}{X - \alpha_k} \pmod{P} .$$

On remarque ensuite que $P/(X - \alpha_k)$ est unitaire de degré $n-1$, pour en déduire que $\ell(P'A) = \sum_{k=1}^{n} A(\alpha_k) = \mathrm{trace}(A)$.

2.2 La base de Horner

Associé au polynôme P, nous avons les *polynômes de Horner* :

$$H_0 = 1,\ H_1 = X + a_1,\ H_2 = X^2 + a_1 X + a_2, \ldots$$
$$\ldots, H_{n-1} = X^{n-1} + a_1 X^{n-2} + \cdots + a_{n-2} X + a_{n-1} .$$

On peut compléter avec $H_n = P$, et on a la relation de récurrence $H_{i+1} = XH_i + a_{i+1}$. Les polynômes de Horner sont surtout connus parce que cette

relation de récurrence fournit (pour un polynôme non nécessairement unitaire de degré n) un schéma d'évaluation avec n multiplications seulement (on peut voir Knuth, Seminumerical algorithms, p. 467). D'après la propriété des degrés étagés, il est clair que $(H_{n-1}, \ldots, H_1, H_0)$ est une base de $\mathbb{R}[X]/P$, que nous appellerons base de Horner.

Proposition 2 *La forme quadratique $Q(P, u)$ de la section 1.2 est l'expression de T dans la base de Horner.*

Les propositions 1 et 2 nous montrent que les formes quadratiques que l'on trouve dans les deux exposés sont en fait la même, à savoir la forme T. Nous allons montrer la proposition 2 par une suite de lemmes.

Lemme 2 *On a :*

$$\frac{P(X) - P(Y)}{X - Y} = \sum_{i=0}^{n-1} H_{n-1-i}(Y)\, X^i \ .$$

Pour montrer ce lemme, on peut montrer

$$\frac{H_p(X) - H_p(Y)}{X - Y} = \sum_{i=0}^{p-1} H_{p-1-i}(Y)\, X^i$$

par récurrence sur p, en utilisant la relation de récurrence entre polynômes de Horner.

Lemme 3 *Pour tout entier i avec $0 \leq i \leq n-1$, on a $\ell(X^i H_{n-1-j}) = \delta_{i,j}$ (de Kronecker).*

Ce lemme se vérifie sans difficulté. Il montre que la forme quadratique $A \mapsto \ell(A^2)$ est non dégénérée sur $\mathbb{R}[X]/P$ et que les bases des monômes et de Horner sont duales pour cette forme. On déduit du lemme 3 :

Lemme 4 *Pour tout élément A de $k[X]/P$, on a $A = \sum_{j=0}^{n-1} \ell(A H_{n-1-j})\, X^j$.*

Revenons maintenant au polynôme $L(P, X, Y)$ de la section 1.2. Il peut s'écrire

$$\begin{aligned}
L(P,X,Y) &= P'(Y)\,\frac{P(X)-P(Y)}{X-Y} - P(Y)\,\frac{P'(X)-P'(Y)}{X-Y} \\
&\equiv P'(Y)\,\frac{P(X)-P(Y)}{X-Y} \pmod{P(Y)} \\
&\equiv P'(Y)\left(\sum_{i=0}^{n-1} H_{n-1-i}(Y)\, X^i\right) \text{ grâce au lemme 2.} \\
&\equiv \sum_{i,j=0}^{n-1} \ell(P' H_{n-1-i} H_{n-1-j})\, Y^j X^i \text{ grâce au lemme 4.}
\end{aligned}$$

Puisque la somme figurant à la dernière ligne est de degré $< n$ en Y, on a en fait

$$L(P,X,Y) = \sum_{i,j=0}^{n-1} \ell(P' H_{n-1-i} H_{n-1-j})\, X^i Y^j \ ,$$

et donc, d'après le lemme 1

$$L(P,X,Y) = \sum_{i,j=0}^{n-1} \mathrm{trace}(H_{n-1-i}H_{n-1-j})\, X^i Y^j \ .$$

Ceci achève la démonstration de la proposition 2.

Ce qu'on a fait dans toute la section 2 peut se faire en remplaçant $\mathbb{R}$ par n'importe quel corps (et $\mathbb{C}$ par un corps de décomposition de P). Le fait qu'on soit sur $\mathbb{R}$ n'intervient qu'à partir du moment où l'on parle de signature de forme quadratique.

3 Pratique

3.1 Trouver la signature

On dispose de plusieurs méthodes pour trouver la signature (s,t) d'une forme quadratique réelle de matrice B de taille n.

1. L'algorithme de Gauss de décomposition en carrés. Par définition de la signature, s est le nombre de carrés avec coefficients positifs et t le nombre de carrés avec coefficients négatifs.
2. Utilisation des mineurs principaux de la matrice B (voir Gantmacher, Théorie des matrices, Dunod, tome 1, p. 305). Le i-ème mineur principal Δ_i est le mineur $i \times i$ construit sur les i premières lignes et colonnes de B. On peut poser $\Delta_0 = 1$. Supposons qu'aucun Δ_i ne soit nul (ceci implique en particulier que la forme quadratique est non dégénérée). Alors s (resp. t) est le nombre d'indices i, avec $1 \leq i \leq n$, tels que $\Delta_{i-1}\Delta_i > 0$ (resp. $\Delta_{i-1}\Delta_i < 0$). Ceci s'explique de la manière suivante. Sous l'hypothèse que tous les Δ_i sont différents de 0, on peut effectuer l'orthogonalisation de Gram-Schmidt sur la matrice B. Ceci signifie qu'il existe une matrice triangulaire supérieure P, avec uniquement des 1 sur la diagonale, telle que $D = {}^tPBP$ soit une matrice diagonale. Si $d_1, \ldots, d_n$ sont les coefficients diagonaux de D, on a $d_i = \Delta_i/\Delta_{i-1}$. Puisque B et D ont même signature, on conclut aisément.

   La matrice H de la section 1.1 n'est pas une matrice symétrique quelconque : ses coefficients ne dépendent que de la somme des indices $i+j$. C'est ce qu'on appelle une *matrice de Hankel*. On peut trouver dans Gantmacher, tome 1, p. 345 une recette spéciale pour calculer la signature de la matrice de Hankel H à partir des signes de ses mineurs principaux, même quand certains de ceux-ci s'annulent.
3. Utilisation de la règle de Descartes (voir Mignotte, Mathématiques pour le calcul formel, PUF, p. 208). Cette règle dit que le nombre de racines strictement positives (comptées avec multiplicité) d'un polynôme réel est inférieur ou égal au nombre de changements de signe dans la suite de ses coefficients. Appliquons cette règle au polynôme caractéristique P_B de la matrice B. On sait que P_B a toutes ses racines réelles. Soit m l'ordre de P_B (la multiplicitéde 0 comme racine). D'après Descartes, s (resp. t) est inférieur ou égal au nombre v_+ (resp. v_-) de changements de signe dans la suite des coefficients de P_B (resp. $P_B(-X)$). On constate que $v_+ + v_- + m$

est inférieur ou égal à n (il faut réfléchir à ce qui arrive quand il y a des coefficients nuls – voir Mignotte, Mathématiques pour le calcul formel, p. 209). On en déduit que $s = v_+$ et $t = v_-$.

3.2 Un exemple en Maple

Voici une feuille de travail Maple avec des procédures qui calculent le nombre de racines réelles d'un polynôme en calculant la signature de sa matrice de Bezout. Les temps de calcul sur un polynôme au hasard de degré 20 que l'on peut voir à la fin de la feuille montrent que la fonction "realroot" de Maple, qui donne en plus des intervalles d'isolation des racines, est beaucoup plus efficace que l'algorithme programmé.

```
> restart: with(linalg):
```

```
Warning, the protected names norm and trace have been redefined and
unprotected
```

```
> Bezout := proc (P,X)
> description "calcule la matrice de Bezout d'un polynôme":
> local d, DP, L:
> d:=degree(P,X): DP:=diff(P,X):
> L:=simplify((P*subs(X=Y,DP)-subs(X=Y,P)*DP)/(X-Y)):
> matrix(d,d,(i,j) -> coeff(coeff(L,X,i-1),Y,j-1))
> end proc:
> variation := proc (P, T)
> local L, LL, t, tt, PP, var;
> description "calcule le nombre de changements de signes dans
> la suite des coefficients";
> L := lcoeff(collect(P,T),T,'t'); PP := P-L*t; var := 0;
> while not PP = 0 do
> LL := lcoeff(collect(PP,T),T,'tt');
> if not sign(L) = sign(LL) then var := var+1
> end if;
> PP := PP-LL*tt : L:=LL
> end do;
> var
> end proc:

> signature := proc( Q )
> description "calcule la signature d'une matrice symétrique";
> local CP:
> CP:=charpoly(Q,T):
> [variation(CP,T),variation(subs(T=-T,CP),T)]
> end proc:

> nbrr := proc( P, X)
> local S;
> description "calcule le nombre de racines réelles de P";
> S:=signature(Bezout(P,X));
> S[1]-S[2]
> end proc:
> P:=(x-2)*(x-1)*(x-3)^4*(x+4)*(x^2+1)*(x^4+1);
```

$$P := (x-2)\,(x-1)\,(x-3)^4\,(x+4)\,(x^2+1)\,(x^4+1)$$

```
> signature(Bezout(P,x));
```

$$[7,\ 3]$$

```
> nbrr(P,x); nbrr(P+1/100,x); nbrr(P-1/100,x);
```

$$4$$
$$3$$
$$5$$

```
> Q:=randpoly(x, dense, degree=20);
```

$$Q := -85\,x^{20} - 55\,x^{19} - 37\,x^{18} - 35\,x^{17} + 97\,x^{16} + 50\,x^{15} + 79\,x^{14} + 56\,x^{13} + 49\,x^{12} + 63\,x^{11} + 57\,x^{10} - 59\,x^{9} + 45\,x^{8} - 8\,x^{7} - 93\,x^{6} + 92\,x^{5} + 43\,x^{4} - 62\,x^{3} + 77\,x^{2} + 66\,x + 54$$

```
> showtime():

> nbrr(Q,x);
```

$$2$$

```
time = 4.10, bytes = 16369770
> realroot(Q);
```

$$[[0,\ 2],\ [-2,\ 0]]$$

```
time = 0.65, bytes = 178182
```

4 Compléments

Une autre méthode pour calculer le nombre de racines réelles d'un polynôme est plus connue : celle donnée par le théorème de Sturm (voir par exemple Mignotte, Mathématiques pour le calcul formel, p. 203).

Il y a d'autres résultats liant racines et signature, que l'on peut trouver dans les deux références Dieudonné et Gantmacher : nombre de racines complexes dans un demi-plan grâce à la signature d'une forme hermitienne (Dieudonné p. 62), calcul de l'indice de Cauchy d'une fraction rationnelle (différence entre le nombre de sauts de $-\infty$ à $+\infty$ et le nombre de sauts de $+\infty$ à $-\infty$) par la signature d'une forme quadratique (Gantmacher, p. 207), démonstration du critère de Routh-Hurwitz pour que toutes les racines d'un polynôme réel aient leurs parties réelles strictement négatives (Gantmacher p. 213). Ce dernier critère est important pour les problèmes de stabilité des systèmes.