Coefficient bounds on the integer characteristic and minimal polynomials

Jean-Guillaume Dumas

▶ **To cite this version:**

Jean-Guillaume Dumas. Coefficient bounds on the integer characteristic and minimal polynomials. 2006. hal-00086820v3

HAL Id: hal-00086820

https://hal.science/hal-00086820v3

Preprint submitted on 21 Jul 2006 (v3), last revised 21 Sep 2007 (v7)

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

ccsd-00086820, version 3 - 21 Jul 2006

Coefficient bounds on the integer characteristic and minimal polynomials

Jean-Guillaume Dumas

July 21, 2006

Abstract

In order to perform deterministic chinese remaindering of the characteristic or minimal polynomial, precise bounds on the size of their integer coefficients and ways to refine these bounds on the fly are presented.

1 Bound on the minors

1.1 Hadamard's bound on the minors

The first bound of the characteristic polynomial coefficient uses Hadamard's bound [7, Theorem 16.6] to show that any integer coefficient of the characteristic polynomial has the order of n bits:

Lemma 1.1. *Let $A \in \mathbb{Z}^{n\times n}$, with $n \geq 5$, whose coefficients are bounded in absolute value by $B > 1$. The coefficients of the characteristic polynomial of A are denoted by c_j. Then*

$$\log_2(|c_j|) \leq \frac{n}{2}\left(\log_2(n) + \log_2(B^2) + 0.21163175\right)$$

Proof. c_j, the j-th coefficient of the characteristic polynomial, is an alternate sum of all the $(n-j)\times(n-j)$ diagonal minors of A. It is therefore bounded by $H(n,j) = \binom{n}{j}\sqrt{(n-j)B^2}^{(n-j)}$. First note, that from the symmetry of the binomial coefficients we only need to explore the $\lfloor n/2\rfloor$ first ones, since $\sqrt{(n-j)B^2}^{(n-j)} > \sqrt{jB^2}^{j}$ for $i < \lfloor n/2\rfloor$. Now, the lemma claims that actually the maximal value must occur within the $\mathcal{O}(\sqrt{n})$ first ones. The lemma is true for $j = 0$ by Hadamard's bound. And for $j = 1$, we have $\log_2\left(H(n,j)\right) < \frac{n}{2}(\log_2(n) + \log_2(B^2) + 0.21163175)$ as soon as $n > 4$, since the difference is decreasing in n. Then from Stirling's formula ($n! = (1+\epsilon(n))\sqrt{2\pi n}\frac{n^n}{e^n}$), we have $\forall i \geq 2\ \binom{n}{i} < \frac{1+\epsilon(n)}{\sqrt{2\pi}}\sqrt{\frac{n}{i(n-i)}}\left(\frac{n}{i}\right)^i\left(\frac{n}{n-i}\right)^{n-i}$. Now first $\frac{1}{12n} < \epsilon(n) < \frac{1}{12n+1}$. Therefore for $n > 4$, $\log_2\left(\frac{1+\epsilon(n)}{\sqrt{2\pi}}\right) \leq -1.296$. Then $\frac{n}{i(n-i)}$ is decreasing in i for $i < \lfloor n/2\rfloor$ so that its maximum is $\frac{n}{2(n-2)}$.

Consider now $K(n,j) = \left(\frac{n}{j}\right)^j \left(\frac{n}{n-j}\right)^{n-j} \sqrt{(n-j)B^2}^{(n-j)}$. We have $\log_2(K(n,j)) = \frac{n-j}{2}\log_2(B^2) + \frac{n}{2}\log_2(n) + \frac{n}{2}T(n,j)$, where $T(n,j) = \log_2(\frac{n}{n-j}) + \frac{j}{n}\log_2(\frac{n-j}{j^2})$. Well $T(n,j)$ is maximal for $j = \frac{-1+\sqrt{1+4en}}{2e}$. We end with the fact that $T(n,j) - \frac{2}{n}1.296 + \frac{1}{n}\log_2(\frac{n}{2(n-2)})$ is maximal over $\mathbb{Z}$ for $n = 15$ where it is lower than 0.208935. The latter is lower than 0.21163175. □

Well,

$$\begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & -1 & -1 & -1 \\ 1 & -1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 & -1 \\ 1 & -1 & -1 & -1 & 1 \end{bmatrix}$$

has $X^5 - 5X^4 + 40X^2 - 80X + 48$ for characteristic polynomial and $80 = \binom{5}{1}\sqrt{4}^4$ is greater than Hadamard's bound 55.9, and less than our bound 80.66661.

Note that this numerical bound improves the one used in [8, lemma 2.1] since $0.21163175 < 2 + \log_2(e) \approx 3.4427$.

1.2 Locating the largest coefficient

Lemma 1.2. *Let $A \in \mathbb{Z}^{n\times n}$, with $n \geq 4$, whose coefficients are bounded in absolute value by $B > 1$. The coefficients of the characteristic polynomial of A are denoted by c_j. Then*

$$\forall j \ |c_j| \leq max_{i=0..\frac{-1+\sqrt{1+4eB^2n}}{2eB^2}} \binom{n}{i} \sqrt{(n-i)B^2}^{(n-i)}.$$

Moreover, the cost of computing the associated bound on the size is

$$\mathcal{O}\left(\frac{\sqrt{n}}{B}\right).$$

The proof is similar to that of lemma 1.1, except that we cannot anymore bound expressions as we go. Indeed say that all values of a function $G(n,j)$ are below one of the values of a function $F(n,j)$ which happend to be e.g. $F(n,2)$, this does not prove that the maximum of G is at $j = 2$. *Proof.* Consider $H(n,j) = \binom{n}{j}\sqrt{(n-j)B^2}^{(n-j)} = B^n G(n,j) B^{-j}$, for $j = 2..\lfloor \frac{n}{2} \rfloor$. Then using Stirling's approximation, we have $G(n,j) = \frac{1+\epsilon(n)}{\sqrt{2p\,i}} \frac{1}{(1+\epsilon(j))(1+\epsilon(n-j))} F(n,j)$. We now call $D(n,j) = \frac{2}{n}\log(F(n,j)) = \left(\log(\frac{n}{n-j}) + \frac{j}{n}\log(\frac{n-j}{j^2}) + \frac{1}{n}\log(\frac{n}{j(n-j)})\right)$. Differentiating D with respect to j yields that the solution of $\ln(\frac{n-j}{j^2}) = 1 + \frac{n-2j}{j(n-j)}$ gives an extremum of $D(n,j)$.

Now we multiply back $F(n,j)$ by both B^{-j} and $\frac{1}{(1+\epsilon(j))(1+\epsilon(n-j))}$ and consider then $Dt(n,j) = D(n,j) - \frac{j}{n}\log(B^2) - \frac{2}{n}\log((1+\epsilon(j))(1+\epsilon(n-j)))$. We approximate the latter with the taylor expansion of ϵ given e.g. in [1] ($\epsilon(i) \approx 1/12/i + 1/288/i^2 - 139/51840/i^3 \ldots$) and differentiate again with respect to j. This yields that the solution $j*$ of $\ln(\frac{n-j}{j^2}) = 1 + \ln(B^2) - \frac{n-2j}{6j^2(n-j)^2}(6j^2 - 6nj - n) + O(\frac{1}{(i^3)(n-i)^2})$ gives an extremum of $D(n,j)$ Then, the roots of

$(6j^2 - 6nj - n)$ are respectively greater than $n/2$ and strictly negative. Thus $\frac{n-2j}{6j^2(n-j)^2}(6j^2 - 6nj - n)$ is always negative and we have $\ln(\frac{n-j*}{j*^2}) \geq 1 + \ln(B^2)$. This proves that

$$j* \leq \frac{-1 + \sqrt{1 + 4eB^2n}}{2eB^2}.$$

Then, for $n > 19$ and $B = 1$ the derivative is negative at $j = 2$ and positive at $j = n/2$ and thus, $j*$ is a maximum. For smaller n, the derivative is always negative and for larger B, the switch in n is larger than 19.

Now for the complexity, we fist compute $\log(H(n, 0)) = \frac{n}{2}\log(nB^2)$ and then $\log(H(n, i + 1)) = \log(H(n, i)) + \log(\frac{n-i}{i+1}) + \frac{n-i-1}{2}\log(n - i - 1) - \frac{n-i}{2}\log(n - i) - \log(B)$. □

2 Eigenvalue bounds

2.1 Minimal polynomial

For the minimal polynomial the Hadamard bound may also be used, but is too pessimistic an estimate, in particular when the degree is small. Therefore, one can use a bound determined by consideration of Gershgörin disks and ovals of Cassini. This bound is of the form β^d where β is a bound on the eigenvalues and d is the degree of the minimal polynomial.

We can then use the following lemma to bound the coefficients of the minimal polynomial:

Lemma 2.1. *Let $B \in \mathbb{C}^{n\times n}$ with its spectral radius bounded by β. Let* $\text{minpoly}_B(X) = \sum_{k=0}^{d} m_i X^i$. *Then $\forall i \in [\![0, d]\!]$, $|m_i| \leq \max\{\sqrt{d\beta}; \beta\}^d$.*

Proof. It suffices to note that $|m_i| \leq \binom{d}{i}\beta^i$ [10, Theorem IV.§4.1], and then bound each one of these with either β^d when $\beta \geq d$ or $d^{\frac{d}{2}}\beta^{\frac{d}{2}}$ when $\beta \leq d$. □

For matrices of constant size entries, both β and d are $\mathcal{O}(n)$. However, when d and/or β is small relative to n (especially d) this may be a striking improvement over the Hadamard bound since the length of latter would be of order $n\log(n)$ rather than $d\log(\beta)$.

This is the case e.g. for the Homology matrices in the experiments of [6]. Indeed, for those, AA^t, the Wishart matrix of A, has very small minimal polynomial degree and has some other useful properties which limit β (e.g. the matrix AA^t is diagonally dominant).

There remains to compute precise bounds on the eigenvalues.

2.2 Ovals of Cassini for the Wishart matrix

The i-th Gershgörin disk is centered at $a_{i,i}$ and has for a radius the sum of the absolute values of the other entries on the i-th row. Gershgörin's theorem is that all of the eigenvalues are contained in the union of the Gershgörin disks [2, 11, 9]. One can then go further and consider the ovals of Cassini [3, 4, 12],

which may produce sharper bounds. Each Cassini oval has two centers q_1, q_2 (two diagonal elements) and two radii r_1, r_2 (e.g. the associated row radii). Then any point λ of this oval satisfies the following: $|\lambda - q_1||\lambda - q_2| \leq r_1 r_2$. Moreover the eigenvalues lie in the union of the ovals [3, Theorem 1].

While computing the Cassini ovals of A is straightforward, computing those of AA^t, whithout performing the expensive matrix product, is not. We remark that it is expensive to compute the bound mentioned above while staying strictly in the black box model, where only matrix-vector products are available [5]. It seems to require two matrix vector products (with A) to extract each row or column of B. But, if one has access to the elements of A, a bound for the spectral radius of B can easily be obtained with very few arbitrary precision operations:

Algorithm 2.1 OCB [Ovals-of-Cassini-Bound]

Require: a matrix $A \in \mathbb{C}^{m \times n}$.
Ensure: $\beta \in \mathbb{R}$, such that for every eigenvalue λ of AA^t, $|\lambda| \leq \beta$.
{Centers}
for all $i = 1$ to m **do**
  $q_i = \sum_{1 \leq j \leq n} a_{ij}^2$
end for
{Radii}
Form $|A|$, the matrix whose entries are the absolute values of those of A.
$v = |A||A|^T[1, 1, \ldots, 1]^T$
for all $i = 1$ to m **do**
  $r_i = v_i - |q_i|$
end for
{Gershgörin bound}
$q = \max_{1 \leq i \leq m} |q_i|$
i_1 such that $r_{i_1} = \max_{i \in [\![1,m]\!]} r_i$
{Cassini bound}
i_2 such that $r_{i_2} = \max_{i \in [\![1,m]\!] \setminus \{i_1\}} r_i$
$\beta = q + \sqrt{r_{i_1} r_{i_2}}$

For a matrix $A \in \mathbb{C}^{m \times n}$ let $\Omega = \max\{m; n;$ number of nonzero elements in $A\}$. Then 2Ω bounds the number of field operations for the matrix vector product, Ax, and for a vector inner product, $x^T x$.

Theorem 2.2. *Let $A \in \mathbb{C}^{m \times n}$ with Ω as described above. Algorithm* Ovals-of-Cassini-Bound *correctly computes a bound on the eigenvalues of AA^t, using no more than 7Ω field operations and $3m$ comparisons.*

Proof. For the correctness of the bound we use the fact that the eigenvalues lie in the union of the ovals of Cassini. Now suppose that q_1, q_2, r_1, r_2 are the two centers and two radii of such an oval. Then any point λ of this oval satisfies the following: $|\lambda - q_1||\lambda - q_2| \leq r_1 r_2$ [3, Theorem 1]. We want to know the maximal

absolute value of such a λ. First, if $|\lambda - q_2| \leq |\lambda - q_1|$, then $|\lambda - q_2| \leq \sqrt{r_1 r_2}$, as $|\lambda| - |q_2| \leq |\lambda - q_2|$, we conclude by $|\lambda| \leq |q_2| + \sqrt{r_1 r_2}$. Replacing q_2 by q_1, the second case is analogous. Therefore β as in the algorithm matches the requirements. The complexity analysis is straightforward. □

References

[1] C. M. Bender and S. A. Orszag. *Advanced Mathematical Methods for Scientists and Engineers*. McGraw-Hill, New York, NY, USA, 1978.

[2] Alfred Brauer. Limits for the characteristic roots of a matrix. I. *Duke Mathematical Journal*, 13:387–395, 1946.

[3] Alfred Brauer. Limits for the characteristic roots of a matrix. II. *Duke Mathematical Journal*, 14:21–26, 1947.

[4] Richard A. Brualdi and Stephen Mellendorf. Regions in the complex plane containing the eigenvalues of a matrix. *American Mathematical Monthly*, 101(10):975–985, December 1994.

[5] Jean-Guillaume Dumas, Thierry Gautier, Mark Giesbrecht, Pascal Giorgi, Bradford Hovinen, Erich Kaltofen, B. David Saunders, Will J. Turner, and Gilles Villard. LinBox: A generic library for exact linear algebra. In Arjeh M. Cohen, Xiao-Shan Gao, and Nobuki Takayama, editors, *Proceedings of the 2002 International Congress of Mathematical Software, Beijing, China*, pages 40–50. World Scientific Pub, August 2002.

[6] Jean-Guillaume Dumas, B. David Saunders, and Gilles Villard. On efficient sparse integer matrix Smith normal form computations. *Journal of Symbolic Computations*, 32(1/2):71–99, July–August 2001.

[7] Joachim von zur Gathen and Jürgen Gerhard. *Modern Computer Algebra*. Cambridge University Press, New York, NY, USA, 1999.

[8] Mark Giesbrecht and Arne Storjohann. Computing rational forms of integer matrices. *Journal of Symbolic Computation*, 34(3):157–172, September 2002.

[9] Gene H. Golub and Charles F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. The Johns Hopkins University Press, Baltimore, MD, USA, third edition, 1996.

[10] Maurice Mignotte. *Mathématiques pour le calcul formel*. Presses Universitaires Françaises, 1989.

[11] Olga Taussky. Bounds for characteristic roots of matrices. *Duke Mathematical Journal*, 15:1043–1044, 1948.

[12] Richard S. Varga. *Matrix iterative analysis*. Number 27 in Springer series in Computational Mathematics. Springer-Verlag, seconde edition, 2000.