Pré-Publicações do Departamento de Matemática
Universidade de Coimbra
Preprint Number 07–39

GAUSS ELIMINATION WITHOUT PIVOTING FOR POSITIVE SEMIDEFINITE MATRICES AND AN APPLICATION TO SUM OF SQUARES REPRESENTATIONS

CARLA FIDALGO

Abstract: It is shown that Gauss elimination without pivoting is possible for positive semidefinite matrices. While we do not claim the method as numerically the most advisable, it allows to obtain sum of squares (sos) representations in a more direct way and with more theoretical insight, than by the usual text book proposals. The result extends a theorem attributed for definite quadratic forms to Lagrange and Beltrami and is useful as a finishing step in recent algorithms by Powers and Wörmann [PW] and Parillo [PSPP] to write polynomials $p \in \mathbb{R}[x] = \mathbb{R}[x_1, ..., x_n]$ as a sum of squares in $\mathbb{R}[x]$ when such a representation exists.

Keywords: positive semidefinite matrices, LU decomposition, Gauss elimination, diagonalization of quadratic forms, sums of squares.
AMS Subject Classification (2000): 15A23, 15A21, 11E25.

Let $q(x) = x^*Ax$ be a quadratic form based on a positive semidefinite (psd) complex matrix A. Then there exists a unitary matrix Q_1, and a nonnegative diagonal matrix Λ such that $A = Q_1^*\Lambda Q_1$. This allows to define $B = Q_1^*\sqrt{\Lambda}Q_1$, and in turn we get $A = B^*B$. Hence $q(x) = x^*B^*Bx$, i.o.w. the sum of squares of the absolute values of the entries of vector Bx is a sum of squares representation for $q(x)$. Subjecting B to a QR-decomposition, $B = Q_2U$, with Q_2 unitary, U upper triangular, one finally gets $q(x) = x^*U^*Ux$, hence $q(x)$ has even a sos representation $q(x) = \sum_{i=1}^{n} |l_i(x_i, ..., x_n)|^2$, with linear forms l_i. If A happens to be positive definite (i.e. psd and nonsingular) then one can show that $\det A(\{1, ..., l\}, \{1, ..., l\}) \neq 0$ for $l = 1, ..., n$; and then more direct ways to obtain such an U are suggested: unique LDU decomposition (from where by symmetry $U = L^*$ follows). These well known facts can be recalled e.g. from looking at Horn and Johnson [HJ] pages p396c-4, p397c-5, p171c-1, p112c-3, p162c-7 (i.e. p. 162, about 7cm from last textrow). Stewart [St, p140c-7] (working with real matrices) suggests a direct way to the (unique) Cholesky factorization $A = U^*U$ of nonsingular A.

Among a dozen or so monographs on numerical linear algebra consulted, the books by Golub and van Loan [GvL, p.146], and Higham [Hi, p.210],

Received December 12, 2007.

are the only ones that mention the possibility of a Cholesky type decompositions for semidefinite real singular systems; but the Matlab code piece given in [GvL] did not work. Relevant parts of the theorem 10.9 in Higham [loc. cit.] can also be obtained by applying work of Mehrmann on LU-decomposition [M1, p219c-6] and [M2, p181c-7], hence without Cholesky type decomposition. But this latter approach, working for the more general class of V-matrices, again is at heart designed for nonsingular systems, and its adaption to singular ones involves a search for permutation matrices in a computationally inefficient manner.

In contrast, we found that for positive semidefinite real matrices (singular or not) triangularization via Gauss elimination without pivoting is always possible and as we shall see useful: matrices obtained by applying sostools (see example 7 below) are quite often rank deficient.

Let $g(X)$ denote the unique result of complete Gauss elimination without pivoting applied to a square matrix X, whenever such is possible (i.e. leading to an upper triangular matrix), and let $\mathrm{diag(X)}$ be the matrix obtained by putting all non-diagonal entries of X equal to zero. Also let $\mathrm{diag(d_1, ..., d_n)}$ be the diagonal matrix made from $d_1, ..., d_n$. Further recall that the Moore Penrose inverse of a diagonal matrix D is a diagonal matrix $D^\dagger$ that has the same 0es as D but entries $1/a$ wherever D has entry $a \neq 0$; see [HJ, p421c-4].

Theorem 1. *Let A be a positive semidefinite real matrix. Then triangularization of A via Gauss elimination without pivoting is possible. If $U = g(A)$ is the resulting upper triangular matrix, then there holds the following:*
a. $u_i := u_{ii} = 0$ iff the i-th row of U is a zero row.
b. If $i_1 < i_2 < ... < i_k$ are the indices of the nonzero rows of U then
$$u_{i_1} u_{i_2} ... u_{i_l} = \det A(\{i_1, ..., i_l\}) \text{ for } l = 1, ..., k.$$
c. Let $D = diag(U)(= \mathrm{diag(u_1, ..., u_n)})$, and $V = D^\dagger U$. Then a lower triangular real matrix L with unit diagonal can be constructed, such that
$$A = LU = LDL^T = V^T DV = (\sqrt{D}L)^T(\sqrt{D}L) = (\sqrt{D}V)^T(\sqrt{D}V).$$
The entries of matrices L, D, U, V occurring are rational in the entries of A.

Proof. We first prove parts a and b by induction on the size n of A. If $n = 1$ there is nothing to prove. Now assume $n \geq 2$ and the theorem already proved for all positive semidefinite matrices of size $< n$.

Case $a_{11} = 0$. Since A is psd, we have $\det A(\{1, i\}) = a_{11}a_{ii} - a_{1i}^2 \geq 0$, hence $a_{1i} = 0$ for $i = 2, ..., n$. Consequently A has the structure

$$A = \begin{pmatrix} 0 & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{pmatrix},$$

with C positive semidefinite.

By induction assumption Gauss elimination without pivoting for C is possible; and by the mechanics of this process, $g(A) = 0 \oplus g(C)$ with $g(C)$, hence $g(A)$ upper triangular. Let $g(C)$ for convenience be indexed with $i, j = 2, ..., n$. Then $g(A)_{ii} = g(C)_{ii}$, $i = 2, ..., n$. Let $2 \leq i_1 < ... < i_k$ be the set of indices of nonzero rows of $g(C)$. By induction assumption these are precisely the indices of nonzero diagonal elements of $g(C)$ and thus of $g(A)$ and we have

$$\begin{aligned} g(A)_{i_1 i_1} \cdots g(A)_{i_l i_l} &= g(C)_{i_1 i_1} \cdots g(C)_{i_l i_l} \\ &= \det C(\{i_1, ..., i_l\}) \\ &= \det A(\{i_1, ..., i_l\}), \end{aligned}$$

proving the induction step for the case $a_{11} = 0$.

Case $a_{11} \neq 0$. Let C be the $(n-1) \times (n-1)$ submatrix occurring in the lower right corner after applying one step of Gauss elimination without pivoting to A. Again we assume the entries of $C = (c_{ij})$ indexed by $i, j = 2, ..., n$. By [M2, p176c-3...177c3] we have the formulae

$$c_{ij} = (a_{ij} a_{11} - a_{i1} a_{1j})/a_{11}, \qquad \det C(I) = \det A(I \cup \{1\})/a_{11},$$

whenever $\{i, j\} \cup I \subseteq \{2, ..., n\}$.

Writing the first of these formulae for c_{ji} we see that C is symmetric, and by the second formula it follows that all principal minors of C are nonnegative. Hence C is positive semidefinite. Reasoning analogously as before, we get this time

$$g(A) = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & g(C) & \\ 0 & & & \end{pmatrix}. \quad (2)$$

If $2 \leq i_1 < i_2 < ... < i_k$ are the indices of the nonzero rows of $g(C)$, then by induction hypothesis,

$$\begin{aligned} g(A)_{i_1 i_1} \cdots g(A)_{i_l i_l} &= g(C)_{i_1 i_1} \cdots g(C)_{i_l i_l} \\ &= \det C(\{i_1, ..., i_l\}) \\ &= \det A(\{1, i_1, ..., i_l\})/a_{11}. \end{aligned}$$

Multiplying with $a_{11} = g(A)_{11}$, we obtain the claim in this case.

c. Again we proceed by induction on n claiming the existence of a unit diagonal lower triangular matrix $\dot{L}$ such that $U = \dot{L}A$ and $\dot{L}A\dot{L}^T = \mathrm{diag(U)} = \mathrm{D}$ is diagonal. The desired $L = \dot{L}^{-1}$. Define

$$\dot{L}_1 = \begin{cases} \begin{bmatrix} 1 & & & \\ -\frac{a_{21}}{a_{11}} & \ddots & & \\ \vdots & & \ddots & \\ -\frac{a_{n1}}{a_{11}} & & & 1 \end{bmatrix} & \text{if } a_{11} \neq 0 \\ \\ I_n & \text{if } a_{11} = 0 \end{cases}$$

Then $\dot{L}_1$ is lower triangular with unit diagonal and we get

$$\dot{L}_1 A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{pmatrix}, \quad \text{and} \quad \dot{L}_1 A \dot{L}_1^T = \begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ 0 & & & \\ \vdots & & C & \\ 0 & & & \end{pmatrix}.$$

In particular $\dot{L}_1 A$ is the result of the first step in the Gauss elimination of A. By induction assumption there exists a real lower triangular matrix $\dot{L}_2$ with unit diagonal such that $\dot{L}_2 C = g(C)$ and $\dot{L}_2 C \dot{L}_2^T = \mathrm{diag(g(C))}$. Consequently, by (2) above, $\dot{L} = (1 \oplus \dot{L}_2)\dot{L}_1$ is the matrix desired for which $A = LU = LDL^T$. Next, we have that $DD^\dagger$ is a matrix with 1s exactly in the diagonal positions belonging to $i \in \{i_1, ..., i_k\}$, and 0s elsewhere. So by part a, and invertibility of L, $DL^T = U = DV$, and so $A = LU = LDV = (DL^T)^T V = U^T V = (DV)^T V = V^T DV$. The remaining factorizations given are now obvious. ∎

Remark 2. a. If A is nonsingular, the representation $A = (\sqrt{D}V)^T(\sqrt{D}V)$ is of course the (unique) Cholesky factorization for A; but there can be infinitely many such factorizations for semidefinite matrices. For example, whenever $a^2 + b^2 = 1$, there holds,

$$\begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 \\ a & b \end{pmatrix} \begin{pmatrix} 0 & a \\ 0 & b \end{pmatrix}.$$

b. Here is a code for computing $U = g(A)$ that will work for positive semidefinite rational matrix as long as numerical errors do not accumulate.

```
n=size(A,1);
for k=1:n
if  ~(A(k,k)==0)
 for j=k+1:n
  A(j,:)=A(j,:)-A(j,k)*(A(k,:)/A(k,k));
 end; end; end;
```

An adaption for precise arithmetic software should work always.

c. We would like to emphasize that the rationality observation at the end of theorem 1 is an insight that seems to be harder to see if one uses Cholesky type factorizations - that is approaches via Higham's theorem.

We now use theorem 1 to give an easy computation of a sum of squares representation of a general positive semidefinite quadratic form $x^T Ax$; it is an extension to psd forms of a theorem that according to [BB] goes back to Lagrange and Beltrami for the nonsingular case; [B] contains for this case an indication of proof. Mirsky [M, p371ff] presents Lagrange's reduction in the more general context of possibly indefinite quadratic forms.

Corollary 3. *Continuing the assumptions and notations of the theorem, let* $l_i(x) = \sum_{j=1}^n v_{ij}x_j$, $i = 1, ..., n$. *Then*

$$x^* Ax = \sum_{\nu=1}^{k} u_{i_\nu} l_{i_\nu}(x)^2.$$

Proof. $x^T Ax = x^T V^T DV x$. Now the vector $Vx = [l_1(x), ..., l_n(x)]^T$, and the claim follows easily. ∎

Example 4. The matrix A below reduces by Gauss elimination without pivoting to the matrix U which gives in turn rise to matrices $D, D^\dagger$, and V.

$$A = \begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 3 \end{bmatrix} \quad U = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 2 \end{bmatrix}, \quad \begin{matrix} D = \mathrm{diag}(1, 0, 2), \\ D^\dagger = \mathrm{diag}(1, 0, 1/2) \end{matrix} \quad V = \begin{bmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Consequently $x^* Ax = 1(x + 2y + z)^2 + 0 + 2z^2$, where on the right $[x, y, z] = [x_1, x_2, x_3]$. ∎

Applying the textbook theory as outlined in the introductory paragraph we would come in this case to our result only via a circuitous path: applying the Matlab command series `[V,D]=eig(A)` (yielding V such that $A = VDV^*$) `B=V sqrt(D); [Q,R]=qr(B');` we get

$$R = \begin{bmatrix} i & 2i & i \\ 0 & 0 & 0 \\ 0 & 0 & 1.4142. \end{bmatrix}.$$

Now, with hindsight, Matlab's 1.4142.. should be read as $\sqrt{2}$ and $i = \sqrt{-1}$ eliminated to yield the sos-representation given above.

Applying alternatively `[L,U,P]=lu(A)` one gets matrices L, U, P such that L is lower triangular with 1s and U is upper triangular and $LU = PA$. One expects rightly that `[L,U,P1]=lu(PA)` yields `P1` $= I_3$. Since the LU-decomposition of A and the LU-decomposition for PA in general are not easily related, we suspect that the routines in our version of Matlab© (at least) cannot be used for finding as directly as desirable 'our' U. For numerical reasons and limited applicability, Gauss elimination without pivoting seems not to be implemented.

We show next how our results can be used as a finishing step in the work of algorithms to write multivariate polynomials as sums of squares. Powers and Wörmann [PW], based on work by Choi, Lam and Reznick [CLR] have proposed algorithms to decide, given $f \in \mathbb{R}[x_1, ..., x_n] = \mathbb{R}[x]$, whether $f(x)$ is a sum of squares of polynomials in $\mathbb{R}[x]$. A necessary condition is of course that f be of even degree $2m$, say. For each monomial occurring in f one considers the family of monomials of degree m such that the former can be written as a biproduct (=product of two) of the latter monomials. Then one forms a column vector $\bar{x}$ of all monomials occurring in the families. Then f is a sum of squares if and only if there exists a positive semidefinite real matrix A, the *Gram* matrix of f, such that $f(x) = \bar{x}^T A \bar{x}$. Closer inspection shows that one can assume $A = A_0 + r_1 A_1 + ... + r_l A_l$ with r_i real parameters and the A_i numerically given real symmetric matrices. As the heart of the algorithmic procedure Powers and Wörmann then propose computation of the characteristic polynomial of A, which by symmetry of A can have has only real roots. A sharp form of Descartes' rule of signs [KS, p40] (derivable also from the apparently weaker statement in [BCR, p14]) thus gives necessary and sufficient conditions in order that A has only nonnegative roots. These conditions are polynomial inequalities in the r_i. Experimenting, or

Tarski's (or similar) decision algorithms [BCR, p17] can now be applied to see whether the r_i can be choosen so that A is psd. Parillo has observed that this decision is much more efficiently dealt with in the context of semidefinite programming; see [PPSP]. We give examples for both approaches.

Example 5. We want to find a sos-representation of $f(x,y) = x^4+y^4-x^3y-xy^3$. The monomials of degree≤ 2 that can occur in biproduct representations of the monomials in f are collected in $\bar{x} = [x^2, xy, y^2]^T$. The general form of a Gram matrix for f is given by

$$B = \begin{bmatrix} 1 & -\frac{1}{2} & r \\ -\frac{1}{2} & -2r & -\frac{1}{2} \\ r & -\frac{1}{2} & 1 \end{bmatrix}$$

The characteristic polynomial of B is

$$-\lambda^3 + (-2r+2)\lambda^2 + (r^2+4r-\frac{1}{2})\lambda + (2r^3 - \frac{3}{2}r - \frac{1}{2}).$$

By Descarte's rule of signs the inequalities in

$$\{-2r+2 \geq 0,\ r^2+4r-\frac{1}{2} \leq 0,\ 2r^3 - \frac{3}{2}r - \frac{1}{2} \geq 0\}$$

must be satisfied. One sees now that A is psd if $r = -\frac{1}{2}$. The eigenvalues of A are $\{0, \frac{3}{2}, \frac{3}{2}\}$, so Cholesky decomposition doesn't exist. But Gauss elimination without row interchange yields the result

$$f = (x^2 - \frac{1}{2}xy - \frac{1}{2}y^2)^2 + \frac{3}{4}(xy - y^2)^2;$$

an optimal result since in [CLR] it is shown that every form in two variables and degree 4 positive semidefinite can be written as sum of two squares.

We next correct an error in [PW].

Example 6. Given $f(x,y,z) = x^4 + 2x^2y^2 + x^3z + z^4$, we need to consider the general form of a Gram matrix for f with respect to the ordered basis

$\{x^2, xy, xz, z^2\}$. This matrix is

$$B = \begin{bmatrix} 1 & 0 & \frac{1}{2} & r \\ 0 & 2 & 0 & 0 \\ \frac{1}{2} & 0 & -2r & 0 \\ r & 0 & 0 & 1 \end{bmatrix}.$$

The characteristic polynomial of B is

$$\lambda^4 - (2r-4)\lambda^3 + (\frac{19}{4} - 8r - r^2)\lambda^2 + (-\frac{5}{4} + 10r + 2r^2 - 2r^3)\lambda + (-\frac{1}{2} - 4r + 4r^3).$$

Thus, B is positive definite for $r = -\frac{1}{2}$. In this case there exists Cholesky decomposition of B. When this happens the sos representation achieved doing Gauss elimination without pivoting on the Gram matrix, and the one achieved by the Cholesky decomposition are exactly the same, but the first method is simpler yielding the final result

$$f = (x^2 + \frac{1}{2}xz - \frac{1}{2}z^2)^2 + 2x^2y^2 + \frac{3}{4}(xz + \frac{1}{3}z^2)^2 + \frac{2}{3}z^4,$$

contrary to the claim in [PW, p102c-6] that f is not sos.

Gram matrices of small rank occur also in pleasing examples like the following instance of the arithmetic geometric inequality $H(x,y,z) \geq 0$.

Example 7. Consider the well known Hurwitz form

$$H(x,y,z) = x^6 + y^6 + z^6 - 3x^2y^2z^2,$$

see [R2]. To find a sos representation, we have to consider all 10 monomials of degree 3 in x, y, z. So we use the ordered basis

$$\{z^3, yz^2, y^2z, y^3, xz^2, xyz, xy^2, x^2z, x^2y, x^3\}.$$

Using SOSTOOLS and SeDuMi, see [PPSP] and [Stu], respectively, we get

$$B = \begin{bmatrix} 2 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 2 & 0 & -1 & 0 & 0 & 0 & 0 & -1 & 0 \\ -1 & 0 & 2 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & 2 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 2 & 0 & -1 & 0 & 0 & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & 2 & 0 & 0 & -1 \\ -1 & 0 & -1 & 0 & 0 & 0 & 0 & 2 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 & 0 & 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 0 & -1 & 0 & -1 & 0 & 0 & 2 \end{bmatrix}$$

as a possible positive semidefinite Gram matrix of $2H$. B is singular and has rank six. Doing Gauss elimination without pivoting, we find the following representation of H

$$\begin{aligned} H(x,y,z) &= 3/4(y^2z - x^2z)^2 + 3/4(y^3 - x^2y)^2 + 3/4(xy^2 - x^3)^2 \\ &\quad +(z^3 - 1/2y^2z - 1/2x^2z)^2 + (yz^2 - 1/2y^3 - 1/2x^2y)^2 \\ &\quad +(xz^2 - 1/2xy^2 - 1/2x^3)^2. \end{aligned}$$

Acknowledgements
This paper was written under the supervision of Professor Alexander Kovačec. The author gratefully thanks him for numerous valuable conversations, helpfull comments and suggestions.

References

[BB] E. F. Beckenbach and Bellman, R. Inequalities, Springer 1971.
[BCR] J. Bochnak, M. Coste, M. F. Roy, Real Algebraic Geometry, Springer 1998.
[B] R. Bellman, Introduction to Matrix Analysis, McGraw Hill 1960.
[CLR] M. D. Choi, T.Y. Lam and B. Reznick, Sums of squares of real polynomials, Proceedings of Symposia in Pure Mathematics, 58 (1995), 103-126.
[GvL] G. H. Golub, C. F. van Loan, Matrix Computations, 2nd. ed., John Hopkins, 1989.
[H] N. J. Higham, Accuracy and Stability of Numerical Algorithms, SIAM 1996.
[HJ] R. J. Horn, C.R. Johnson, Matrix analysis, Cambridge, 1985.
[KS] M. Knebusch, C. Scheiderer, Einführung in die relle Algebra, Vieweg, 1989.
[M1] V. Mehrmann, On classes of matrices containing M-matrices and Hermitian positive semidefinite matrices, Linear Algebra Appl. 58: 217-234 (1984).
[M2] V. Mehrmann, On the LU Decomposition of V-Matrices, Linear Algebra Appl. 61: 175-186 (1984).
[M] L. Mirsky, An Introduction to Linear Algebra, Dover (reprint) 1990.

[PPSP] S. Prajna and A. Papachristodoulou and P. Seiler and P. A. Parrilo, SOSTOOLS: Sum of squares optimization toolbox for MATLAB, (2004), available from `http://www.cds.caltech.edu/sostools` and `http://www.mit.edu/~parrilo/sostools`.

[PW] V. Powers and T. Wörmann, An algorithm for sums of squares of real polynomials, Journal of Pure and Applied Algebra, 127 (1998), 99-104.

[R1] B. Reznick, A quantitative version of Hurwitz' theorem on the arithmetic-geometric inequality, J. reine angew. Math., 377 (1987), 108-112.

[R2] B. Reznick, Forms derived from the arithmetic-geometric inequality, Math. Ann., 283 (1989), 431-464.

[St] G. W. Stewart, Introduction to Matrix Computations, Acad. Press 1973.

[Stu] J. Sturm, SeDuMi version 1.05, (2001), available from `http://fewcal.uvt.nl/sturm/software/sedumi.html`.

Carla Fidalgo
Departamento de Física e Matemática, Instituto Superior de Engenharia de Coimbra, Rua Pedro Nunes, Quinta da Nora, 3030-199 Coimbra, Portugal; Department of Mathematics, University of Coimbra, 3001-454 Coimbra, Portugal

E-mail address: `cfidalgo@isec.pt`