On a matrix decomposition

Benoît Meister, LSIIT/ICPS, UMR 7005
Université Louis Pasteur
Strasbourg, France

April 8, 2002

1 Aim of the paper

This short report deals with a decomposition of a square integral matrix M of size n defined by:

$$M = N_1 \cdot U_2 \ldots N_{m-1} \cdot U_m \cdot N_m \cdot U_0$$

with $m \leq n-1$ and such that :

- N_i are square diagonal integral matrices with nonnegative elements, with $N_{i_{k,k}} =$ gcd of the k^{th} column of $M_i = N_1 \cdot U_2 \ldots N_{i-1} \cdot U_i \cdot N_i$
- U_i are integral unimodular matrices and for $i \in [2..m]$, lower triangular with nonnegative diagonal elements.

Let us try to characterize these matrices as precisely as possible. We will further consider matrices as loop transformations, so the revealed properties might help us to build proofs upon program transformations.

In section 2, we prove the existence of such a decomposition for any full row rank matrix. Then, we give an algorithm to compute it in section 3. Both are based on the left Hermite Normal Form (HNF) of a matrix, and on column division by its gcd. After that, we show in section 4 some matrix properties that may be useful to exploit this decomposition in a deeper way. Finally, we show how this decomposition applies more generally to rectangular integral matrices in section 5, and we briefly present our implementation of a C function performing the decomposition in section 6.

2 Existence of the decomposition

We want to prove that such a decomposition exists for any square full rank matrix M. First, there always exist a left HNF for M: $M = H \cdot U_0$, where H is a square lower triangular integral matrix with positive diagonal elements, and U is a unimodular integral matrix. Let $b = max\{k|H_{k,k} > 1\}$.

1.
   - If $\rho = det(H) > 1$: As H is integral with positive diagonal elements, we have $H_{k,k} = 1, \forall k > b$. Since H is a left HNF, a diagonal element is stricly greater than the other elements of its row ([Sch86], section 4.1). Therefore we have $H_{k,i} = 0, \forall k > b, \forall i < k$. Thus, the elements of each row $H_{k,\cdot}$ with $k > b$ are equal to zero except the diagonal element, which equals 1. As H is lower triangular, the only non-zero element of the columns $H_{\cdot,i}$ are the diagonal elements $H_{i,i}, \forall i \geq b$. In particular, the b^{th} column of H has only one non-zero element, whose value is greater than 1.

     Then H can be decomposed into :

     $$H = K \cdot N,$$

     where N is a diagonal integral matrix with

     $$\begin{cases} N_{i,i} = 1 & \text{if } i \neq b \\ N_{b,b} = H_{b,b} \end{cases}$$

     and K is equal to H except that $K_{b,b} = 1$. Note that $det(N) = N_{b,b} > 1$ so $det(K) < det(H)$.
   - If $det(H) = 1$, then H is the identity matrix, since it is a HNF ([Sch86], theorem 4.3).

2. Let K be the lower triangular integral matrix with positive diagonal elements obtained by the decomposition in 1, with the same value for b. $K_{i,b} = \delta_{i,b}$, where δ is the kronecker symbol. As $K_{b,b} = 1$, we can nullify all the elements of the b^{th} row except the diagonal element, by substracting integral multiples of the b^{th} column to the i^{th} columns, $i < b$. Let H' be the resulting matrix : $K = H' \cdot U'$, where U' describes the latter column operations. H' is a square lower triangular matrix with nonnegative elements, equal to H except for the b^{th} row. Moreover, for each row the diagonal element is striclty greater than the other elements. As a consequence H' is a left Hermite

Normal Form of K. Since a Hermite Normal Form is unique, H' is *the* left Hermite Normal Form of K. Thus, H' has the same properties as H, with $max\{k|H'_{k,k} > 1\} < b$.

We see that the decomposition described in 1 and the HNF decomposition can be done one after each other (H' playing the role of H). Applying 1 and 2 to a lower triangular integral matrix with positive diagonal elements H that is the left HNF of a matrix decomposes it into $H = H'.U'.N$, where H' is a lower triangular integral matrix with positive diagonal elements that is a left HNF of a matrix (namely K).

We can generalize the decomposition described in 1. Dividing $H_{.,i}$ by its gcd still gives a lower triangular integral matrix with positive diagonal elements. So we can decompose H in a more direct manner into $H = K \cdot N$, where K is a lower triangular integral matrix with positive diagonal elements and N is a diagonal matrix with $N_{i,i} = gcd(H_{.,i})$. In the following, let us call this operation *multidimensional linear compression*. Note that it may reduce the number of matrices in the decomposition.

Moreover, the first kind of decomposition concerns matrices with a determinant $det(H)$ greater than one, and the determinant of the resulting matrix H' is less than $det(H)$. Besides, the value of $max\{k|H_{k,k} > 1\}$ decreases after applying the first and then the second decomposition. The decreasing value of b and of the determinant, which both are integral and have a lower bound, proves that the number of steps in the decomposition is finite. So is the number of matrices in the decomposition, too.

The matrix decomposition proposed here consists in decomposing H by the way described in 1: $H = K \cdot N$, then decomposing the lower triangular integral matrix with positive diagonal elements K as in 2: $K = H' \cdot U'$. Obviously, it is the principle of the algorithm we present in the next section.

3 Algorithm

The presented decomposition is actually made of two kinds of decomposition. One of the two is the left Hermite normal form, which is well-known, and effectively implemented as well, notably in the Polylib [Wil93]. Here, it will be called *left_HNF*. The second one will be called *Compression*, as the matrix N can be see as a linear compression matrix.

```
Compression(input : a square  lower triangular integral matrix with
positive diagonal elements H of rank n)
```

```
// first, initialize matrix N to null, and K = H .
for i = 1 to n
    N[i,i] = gcd(H[.,i])
    for j = i to n
         K[j,i] = H[j,i] / N[i,i]
    enfor
endfor
→ output: K.N
```

```
Dec(input: a lower triangular integral matrix with positive diagonal
elements H of rank n)

K.N = Compression(H)
H'.U' = left_HNF(K)
→ output: if H = identitiy matrix: H;
          else: Dec(H').U'.N
```

```
Decomposition( input: a matrix M of full row rank n)

H.U = left_HNF(M)
→ output : Dec(H) · U
```

The wanted decomposition of a matrix M is thus : Decomposition(M).

4 Some matrix properties

The purpose of this section is to exhibit some additional properties of the matrix in the decomposition. In particular, the matrices $U_i, \forall i \in [2..m]$ are unit lower triangular, i.e. lower triangular with diagonal elements equal to 1. We denote by $\mathcal{A}$ the set of unit lower triangular matrices.

4.1 Stability of $\mathcal{A}$ regarding product and inversion

Stability of $\mathcal{A}$ regarding matrix product is trivial, so it won't be proven here. The identity matrix $\mathbb{I}$ belongs to $\mathcal{A}$. Let $M \in \mathcal{A}$. We have $M \cdot M^{-1} = \mathbb{I}$. M is a lower triangular matrix, so M^{-1} is a lower trianguler matrix as well. Moreover, expliciting the product gives us: $M_{k,k} \cdot M^{-1}_{k,k} = 1, k \in [1..n]$, which

is equivalent to saying $M^{-1}_{k,k} = 1$. Thus, $M^{-1} \in \mathcal{A} : \mathcal{A}$ is stable regarding matrix inversion.

4.2 Matrices U_i are unit lower triangular $\forall i \in [1..m]$

The left Hermite Normal Form K of a square matrix $H = K \cdot U$ is lower triangular. A matrix $U_i, \forall i \in [2..m]$ is obtained from the Hermite Normal Form of a lower triangular integral matrix with positive diagonal elements (H), which produces another lower triangular integral matrix with positive diagonal elements (K). Hence, stability of the set of lower triangular matrices with positive diagonal elements by matrix product and inversion tells that U_i is a lower triangular matrix. More precisely, we have $U = K \cdot H^{-1}$, (H is invertible). Since the set of lower triangular matrices with positive diagonal elements is stable by matrix inversion, H^{-1} is in this set. Then, stability of the set of lower triangular matrices with positive diagonal elements by product gives us: U is in this set. Moreover, U is unimodular then its diagonal elements are equal to 1.

Actually, as the coefficients of K are always lesser than the corresponding coefficients of H, $H_{k,j}$ can be expressed as a positive linear combination of elements of $K_{k,..}$ These positive coefficients form the k-th row-vectors of the corresponding matrix U_k. Hence, U_k is a unit lower triangular matrix with nonnegative elements. Unfortunately, the set of unit triangular matrices with nonnegative elements is not stable regarding inversion, so considering matrix U_k as an element of this set may not help a lot.

5 Extensions

5.1 Extension to non-square matrices

There always exist a left Hermite Normal Form for any full row rank rectangular matrix M: $M = [H\ 0].U$, where H is a square lower triangular integral matrix (and U is integral unimodular). The proposed decomposition process can be applied to H, to obtain the decomposition M. The null columns of $[H\ 0]$ can be ignored, as they only produce additional null columns.

Moreover, there always exists a right Hermite Normal Form for any full column rank rectangular matrix M: $M = U \cdot \begin{bmatrix} H \\ 0 \end{bmatrix}$, where H is a square upper triangular integral matrix. Symetrically, a similar decomposition process can be applied to H to obtain a decomposition, if we operate on rows instead of columns. The corresponding decompositions will be :

- linear compression over rows instead of columns: $H = N.K$
- the right HNF instead of the left HNF: $K = U.H'$ where H' and U are integral upper triangular and U is a unit upper triangular matrix.

The overall form of the decomposition is then:

$$M = U_0 \cdot N_m \cdot U_m \cdot N_{m-1} \ldots U_2 \cdot N_1$$

with $m \leq n$.

5.2 A particular form of the decomposition

We can force the decomposition to process one dimension of the matrix M after each other, giving a decomposition with $m = n$. The only difference with the formerly presented decomposition is the possible introduction of matrices U_i and N_i equal to the identity matrix. Each matrix $U_k, \forall k \in [1..n]$ of the decomposition is equal to the identity matrix except for the k^{th} row, which has nonnegative elements at the left of the diagonal. Each matrix $N_k, \forall k \in [1..n]$ is equal to the identity matrix except for the k^{th} diagonal element which is greater than 1. Then, N_k is characterized by its element $\alpha_k = N_{k,k}$. Actually, we can describe a matrix M by :

- its matrix U_0
- the ordered set of vectors $v_k = U_{k_{k,.}}, \forall k \in [0..n]$
- the n-vector of positive integers $\alpha_k, \forall k \in [1..n]$. Note that, as a result of the general form of $U_i, \forall i \in [1..n]$, the k^{th} row of the matrix $U = \prod_{k=1}^{n} U_k$ is equal to the k^{th} row of U_k.

It might be interesting to consider the whole set of matrices with the same U (and maybe U_0) but with a distinct α_k as a *family* of matrices. Also, it might be interesting to study the unimodular matrix $U \cdot U_0$, common to all the matrices of such a *family*. This might be a part of my future works.

6 Implementation

The decomposition presented in this article has been implemented in C for full row rank matrices using polylib [Wil93]. The function takes a full row rank matrix as its input, and returns the list of matrices of the decomposition. The user can choose which kind of decomposition is returned, allowing or not multidimensional compression. It is currently available at the address *http://icps.u-strasbg.fr/ meister/decomp.c.tgz* .

References

[Sch86] A. Schrijver. *Theroy of Linear and Integer Programming.* John Wiley and Sons, New York, 1986. ISBN 0-471-90854-1.

[Wil93] D.K. Wilde. A library for doing polyhedral operations. Technical report, Research Report 785, IRISA, 1993. freely available at http://icps.u-strasbg.fr/Polylib.