Hermite normal form: Computation and applications

February 24, 2009

1 Uniqueness of Hermite normal form

In the last lecture, we showed that if B is a rational matrix of full row rank, then there is a unimodular matrix U such that $\left[H \mid 0\right] = BU$ is a matrix in Hermite normal form. In particular, $\Lambda(B) = \Lambda(H)$, and consequently, every group generated by rational vectors has a basis in Hermite normal form, and therefore, is a lattice. Now, we prove that this basis H is actually unique.

We shall need the following simple observation: If $H = \left[h_1, h_2, \ldots, h_n\right]$ is a matrix in Hermite normal form and $b \in \Lambda(H)$ is a vector with first $i-1$ components equal to zero, then

$$h = \sum_{j=1}^{n} x_j h_j = \sum_{j=i}^{n} x_j h_j$$

for some integers x_i. Indeed, suppose that $i > 1$; since h_1 is the only vector with non-zero first component, we necessarily have $x_1 = 0$, and proceeding by induction we can show that $x_1 = x_2 = \ldots = x_{i-1} = 0$.

Theorem 1 (Uniqueness of Hermite normal form)**.** *Let B be a rational matrix of full row rank. Then there is a unique matrix H in Hermite normal form such that $\Lambda(H) = \Lambda(B)$.*

Proof. Existence of such a matrix H was shown in the previous lecture. For uniqueness, suppose that $H = \left[h_{ij}\right]$ and $H' = \left[h'_{ij}\right]$ are two matrices in Hermite normal form and $\Lambda = \Lambda(H) = \Lambda(H')$ (we drop all-zero columns and assume that H and H' are square non-singular matrices). Yet, suppose that $H \neq H'$ and choose i and j such that $h_{ij} \neq h'_{ij}$ and i is as small as possible; without loss of generality, we assume that $h_{ij} > h'_{ij}$. The j-th column of H, say h_j, and the j-th column of H', say h'_j, both belong to Λ, and therefore, $h_j - h'_j \in \Lambda$. Therefore, $h_j - h'_j$ can be expressed as an integral linear combination of columns of H. By the choice of i, the vector $h_j - h'_j$ has zeros in its first $i-1$ components. Thus, it is actually an integral linear combination of the columns $i, i+1, \ldots$ of H. But among these columns, only the i-th column has a non-zero entry in the i-th component, which implies $h_{ij} - h'_{ij} = z h_{ii}$ for some integer z. But $h_{ij} < h_{ii}$ and $h'_{ij} < h'_{ii} < h_{ii}$, whence $|h_{ij} - h'_{ij}| < h_{ii}$. It follows that z must be zero, which is a contradiction. □

2 Linear Diophantine equations

Hermite normal form appears to be very useful for solving systems of linear Diophantine equations. Let A be a matrix and b a vector, and consider the problem of finding an integral vector x satisfying the system $Ax = b$. In fact, we may assume that A has full row rank; otherwise, we may remove redundant equations from the system. Yet, we assume all input data to be rational.

We can find a unimodular matrix U such that $\left[H \mid 0\right] = AU$ is a matrix in Hermite normal form. Now, we can apply a standard trick to transform a system of linear equations (and actually, a system of linear inequalities, too) into a more suitable form: $Ax = b$ is equivalent to $(AU)(U^{-1}x) = b$, and therefore, $\left[H \mid 0\right]z = b$, where $z = U^{-1}x$ is integral if and only if x is integral. We can observe that the last components of z may take arbitrary values, while feasibility of the system $\left[H \mid 0\right]z = b$, and therefore, of $Ax = b$, depends only on whether the vector $H^{-1}b$ is integral. Now, we can derive the following condition for a system of linear Diophantine equations to have a solution.

Theorem 2. *Let A be a rational matrix of full row rank and let b be a rational vector. Then the system $Ax = b$ has an integral solution x if and only if $y^{\mathrm{T}}b$ is an integer for each rational vector y, for which the matrix $y^{\mathrm{T}}A$ is integral.*

Proof. If x is an integral vector such that $Ax = b$ and y is a rational vector such that $y^{\mathrm{T}}A$ is integral, then $y^{\mathrm{T}}b = (y^{\mathrm{T}}A)x$ is clearly an integer. For the converse, we apply a unimodular transformation to obtain the equivalent system $\left[H \mid 0\right]z = b$, where $\left[H \mid 0\right]$ is a matrix in Hermite normal form, $z = U^{-1}x$ (z is integral if and only if x is integral). Yet, $y^{\mathrm{T}}A$ is integral if and only if $y^{\mathrm{T}}\left[H \mid 0\right] = (y^{\mathrm{T}}A)U$ is integral. Since $H^{-1}\left[H \mid 0\right] = \left[I \mid 0\right]$ is an integral matrix, $H^{-1}b$ must be an integral vector. But then

$$z_0 = \begin{bmatrix} H^{-1}b \\ 0 \end{bmatrix}$$

is an integral solution of the system $Bz = b$. This completes the proof. □

Suppose that there is an integral solution, i.e., the vector

$$z_0 = \begin{bmatrix} H^{-1}b \\ 0 \end{bmatrix},$$

which is the unique solution of the system $Hz = b$, is integral. Then the whole set of solutions has the form

$$\left\{z_0 + \alpha_{m+1}e_{m+1} + \alpha_{m+2}e_{m+2} + \ldots + \alpha_n e_n : \alpha_{m+1}, \alpha_{m+2}, \ldots, \alpha_n \in \mathbb{Z}\right\},$$

where m is the rank of H (since we assumed A to have full row rank, this is just the number of equations in the system), $e_{m+1}, e_{m+2}, \ldots, e_n$ are the unit vectors corresponding to the zero columns in $\left[H \mid 0\right]$. Since $z = U^{-1}x$, U is unimodular, the set of solutions in the original space of x-variables is

$$\left\{x_0 + \alpha_{m+1}x_{m+1} + \alpha_{m+2}x_{m+2} + \ldots + \alpha_n x_n : \alpha_{m+1}, \alpha_{m+2}, \ldots, \alpha_n \in \mathbb{Z}\right\},$$

where $x_0 = Uz_0$ is a solution of the system $Ax = b$ and $x_i = Uz_i$ $(i = m+1, m+2, \ldots, n)$ form a basis of the lattice of solutions of the system $Ax = 0$.

Theorem 3. *Let A be a rational $m \times n$-matrix of full row rank and let b be a rational m-vector. If the system $Ax = b$ has an integral solution, then the whole set of integral solutions has the form*

$$\left\{x_0 + \alpha_{m+1}x_{m+1} + \alpha_{m+2}x_{m+2} + \ldots + \alpha_n x_n : \alpha_{m+1}, \alpha_{m+2}, \ldots, \alpha_n \in \mathbb{Z}\right\},$$

where x_0 is a solution of the system $Ax = b$, x_i's are some linearly independent integral vectors satisfying $Ax = 0$.

3 Polynomial algorithm for Hermite normal form

In the last lecture, we proved that every rational matrix can be brought into Hermite normal form by a sequence of elementary column operations, to which we include

(1) swapping two columns,

(2) multiplying a column by -1, and

(3) adding an integer multiple of a column to another column.

The proof was essentially algorithmic but we have not analyzed its running time yet.

First, we remind the algorithm. Given a $n \times m$-matrix B of full row rank, we construct a sequence of matrices $B_1, B_2, \ldots, B_n$, where B_k has the form

$$B_k = \begin{bmatrix} H_k & 0 \\ C_k & D_k \end{bmatrix},$$

with H_k being a square non-negative lower-triangular matrix of order k, where each row has the unique maximal entry, which is located on the main diagonal. In order to derive the matrix B_{k+1}, we first eliminate all but one non-zero entries in the first row of D_k. For this, we essentially execute the Euclidean algorithm: after multiplying, when needed, columns by -1, we may assume that all entries in the first row of D_k are non-negative. Now, if $d_i > d_j$ are two non-zero entries in the first row of D_k, we add $-\lfloor d_i/d_j \rfloor$ times the j-th column to the i-th column; particularly, we assign $d_i := d_i - \lfloor d_i/d_j \rfloor d_j$. If d_i is still non-zero, we repeat the process (but now $b_j > b_i$, thus we add $-\lfloor d_j/d_i \rfloor$ times the i-th column to the j-th column). From the running time analysis of the Euclidean algorithm, we know that after $O(\text{size}(d_i)\text{size}(d_j))$ iterations, one of the numbers becomes zero. We apply the same procedure to ensure that all but one entries in the first row of D_k became zero and the only non-zero entry, say d, can be located on the main diagonal by swapping the columns. Finally, we need to ensure that d is the largest number in the $(k+1)$-th row: for this purpose, we add appropriate multiples of the $(k+1)$-th column to the other columns (division with remainder).

It is clear that the algorithm would be (weakly) polynomial if we could guarantee that all entries appearing during the execution of the algorithm have polynomial size. How can we do this? First, we need to check that the "answer" (i.e., Hermite normal form of a matrix) itself has a polynomial size. Let B be an integral matrix of full row rank and let $\begin{bmatrix} H \mid 0 \end{bmatrix}$ be its Hermite normal form. Observe that in the above procedure, d is actually the greatest common divisor of the elements in the first row of D_k. This observation can easily be extended by induction to show that if $h_{11}, h_{22}, \ldots, h_{nn}$ are the diagonal entries of H, then their product is the greatest common divisor of all subdeterminants of B of order n. All other entries in i-th row of H are non-negative and smaller than h_{ii}, and hence, also bounded by a polynomial in the size of B. Thus, the size of the Hermite normal form of B is indeed polynomial in the size of B.

Lemma 4. *Let B be a rational matrix of full row rank and let $\begin{bmatrix} H \mid 0 \end{bmatrix}$ be its Hermite normal form. Then the size of H is bounded by a polynomial in the size of B.*

Recall that if Λ' is a sublattice of a lattice Λ, then

$$D\Lambda \subseteq \Lambda' \subseteq \Lambda, \tag{1}$$

where D is the index of Λ' in Λ. We assume that B is an integral matrix (otherwise, we can find the least common multiple of all denominators in B, say δ, and proceed with the matrix δB) with n rows. Let B' be a square non-singular submatrix of B of order n and consider the lattices $\Lambda = \mathbb{Z}^n$ and $\Lambda' = \Lambda(B')$. Clearly, Λ' is a sublattice of Λ and the index of Λ' in Λ is equal to $D = \det(\Lambda') = |\det(B')|$. Then inclusion (1) states that $D\mathbb{Z}^n \subseteq \Lambda(B')$. In other words, adding the columns $De_1, De_2, \ldots, De_n$ to the matrix B does not change the lattice $\Lambda(B)$ (all of these vectors can be expressed as integral linear combinations of columns from B'), and therefore, does not change the Hermite normal form of B, except of adding extra zero columns to it.

Now, we can run our algorithm on the matrix $\left[B \mid DI\right]$. If, at some step of the algorithm, an entry b_{ij} grows to much, we can add an appropriate multiple of the i-th column in the matrix DI to the j-th column in the matrix B, to make b_{ij} lie between 0 and D. Thus, we can keep the size of any entry bounded by $O(\text{size}(D))$, which is polynomial in the input size. Consequently, the whole algorithm runs in polynomial time.

Theorem 5. *The Hermite normal form of a rational matrix B of full row rank can be computed in polynomial time.*

Similarly, we can find a unimodular matrix U such that BU is in Hermite normal form.

Theorem 6. *There is a polynomial algorithm that, provided a rational matrix B of full row rank, computes a unimodular matrix U such that BU is in Hermite normal form.*

Proof. Again, let B' be a square non-singular submatrix of B; without loss of generality, we may assume that B has the form $B = \left[B' \mid B''\right]$. Let $\left[H \mid 0\right]$ be the Hermite normal form of B. We need to find a unimodular matrix U such that $\left[H \mid 0\right] = \left[B' \mid B''\right]U$. If B was a square matrix, computation would be straight-forward. But we can make it square by adding extra rows: it is easy to see that the Hermite normal form of the matrix

$$\begin{bmatrix} B' & B'' \\ 0 & I \end{bmatrix}$$

has the form

$$\begin{bmatrix} H & 0 \\ H' & H'' \end{bmatrix}$$

for some H' and H''. By Lemma 4 the size of the latter matrix is bounded by a polynomial in the size of the first, and therefore, in the size of B. Consequently, the matrix

$$U = \begin{bmatrix} B' & B'' \\ 0 & I \end{bmatrix}^{-1} \begin{bmatrix} H & 0 \\ H' & H'' \end{bmatrix}$$

can be computed in polynomial time. □

4 Algorithm for linear Diophantine equations

It is rather straight-forward now to derive a polynomial algorithm to solve a system of linear Diophantine equations. Let A be a rational matrix and b be a rational vector, and consider the system $Ax = b$. By exploiting the Gaussian elimination algorithm, we can remove—in polynomial time—redundant rows and assume that A has full row rank (of course, if the system $Ax = b$ has no fractional solution, it does not have an integral solution as well). Now, we can compute—in polynomial time—a unimodular matrix U such that $AU = \left[H \mid 0\right]$ is a matrix in Hermite normal form, and construct an equivalent system $\left[H \mid 0\right] y = b$, where $y = U^{-1}b$. It remains to solve the system $Hy = b$, which can be done by Gaussian elimination, and apply the unimodular transformation to transform the solution to the original x-variables: very similar to what we have done in Section 2. This yields the following theorem.

Theorem 7. *A system of linear Diophantine equations $Ax = b$ can be solved in polynomial time.*

We may return the whole set of solutions in the form, as given in Theorem 3.