1 Outline

In this lecture, we study

- Hermite normal form,
- Unimodular matrices,
- Solving the equality constrained integer feasibility problem

2 Solving a system of equations with integer constraints

We consider the problem finding a solution satisfying the following constraints.

$$Ax = b \quad \text{and} \quad x \in \mathbb{Z}^d \tag{$*$}$$

where $A \in \mathbb{Q}^{m\times d}$ and $b \in \mathbb{Q}^m$ consist of rational entries. Remember that integer programming is NP-hard. In fact, the feasibility problem, the problem of finding a solution satisfying an inequality system and integrality constraints

$$Ax \leq b \quad \text{and} \quad x \in \mathbb{Z}^d$$

is NP-hard. However, when the linear constraints are given by all equality constraints as in ($*$), the feasibility problem is in the complexity class P.

Why is this true? The basic intuition is as follows. Suppose that A is a $d \times d$ nonsingular square matrix. Then $Ax = b$ holds if and only if $x = A^{-1}b$. Hence, $x = A^{-1}b$ is the unique solution satisfying the equality system. What this means is that the problem has a (unique) solution if $A^{-1}b \in \mathbb{Z}^d$. If $A^{-1}b \notin \mathbb{Z}^d$, it means that the feasibility problem does not have a solution. In general, the number of rows in A, given by m, can be smaller than A, and the matrix is not necessarily of full row rank. Nevertheless, the same intuition holds for the general case, for which we will develop mathematical tools and an algorithm to find a solution satisfying ($*$) and show that the algorithm runs in polynomial time.

2.1 Hermite normal form

The first step is to transform the equality system $Ax = b$ into something easier to deal with. We reduce the constraint matrix A into the so-called **Hermite normal form** using **unimodular column operations**.

Definition 5.1. A matrix A is in **Hermite normal form** if A is of the form $A = \begin{bmatrix} D & 0 \end{bmatrix}$ where D is a square nonnegative matrix such that

(a) D is lower triangular.

(b) $d_{11} > 0$ and $0 \leq d_{ij} < d_{ii}$ for $i = 2, \ldots, d$ and $j < i$.

In particular, D is nonsingular.

Suppose that A is already in Hermite normal form. Then $(*)$ is easy to solve. Let y and z denote the partial vectors corresponding to D and 0, respectively.

$$Ax = \begin{bmatrix} D & 0 \end{bmatrix} \begin{bmatrix} y \\ z \end{bmatrix} = Dy.$$

Then we get $Dy = b$ from $Ax = b$. Since D is nonsingular, $Dy = b$ if and only if $y = D^{-1}b$. Then $Ax = b$ is equivalent to $x = (D^{-1}b, z)$. If $D^{-1}b \notin \mathbb{Z}^m$, then $(*)$ has no solution. If $D^{-1}b$ is integral, then $x = (D^{-1}b, 0)$ is an integer solution to $(*)$, and in general, $x = (D^{-1}b, z)$ for any integral z gives rise to a solution.

Motivated by this, the strategy is to transform the system $Ax = b$ to another system $Hx = b'$ where H is in Hermite normal form. Here, H would be the Hermite normal form of A. As mentioned above, to deduce H, we apply unimodular column operations that are defined below.

Definition 5.2. There are three kinds of unimodular column operations.

(*i*) Interchanging two columns.

(*ii*) Multiplying a column by -1.

(*iii*) Adding an integer multiple of a column to another column.

Theorem 5.3. *Let A be a rational matrix with **full row rank**. Then A can be transformed into Hermite normal form by a finite sequence of unimodular column operations.*

Proof. If H is the Hermite normal form of A, then the Hermite normal form of kA for some positive scalar k is kH. In particular, we may multiply A by the least common multiple of the denominators of its entries so that the resulting matrix has all entries integral. Thus, we may assume that all entries of A are integers without loss of generality.

Suppose that $A = \begin{bmatrix} D & 0 \\ B & C \end{bmatrix}$ where $\begin{bmatrix} D & 0 \end{bmatrix}$ is in Hermite normal form. At the beginning, $\begin{bmatrix} D & 0 \end{bmatrix}$ may have no row, in which case, $A = \begin{bmatrix} B & C \end{bmatrix}$. Basically, we will enlarge the row submatrix of A that is in Hermite normal form, by transforming the first row of submatrix $\begin{bmatrix} B & C \end{bmatrix}$. **Note that the first row of C must have a nonzero entry. Otherwise, the first row of $\begin{bmatrix} B & C \end{bmatrix}$ is implied by $\begin{bmatrix} D & 0 \end{bmatrix}$, contradicting the assumption that A is of full row rank.**

- Using operations (i) and (ii) in Definition 5.2, we transform C so that $c_{11} \geq c_{12} \geq \cdots \geq c_{1k} \geq 0$. **(Note: As the first row of C must have a nonzero entry, C_{11} has to be strictly positive.)**
- If $c_{1j} > 0$ for some $j \geq 2$, then using operation (iii) in Definition 5.2, we can subtract the jth column of C from the first $j - 1$ columns of C.

This procedure keeps the first row of C still nonnegative while decreasing the value of $c_{11} + \cdots + c_{1k}$. Moreover, $c_{11}, \ldots, c_{1k}$ remain to be integers. We may repeat this procedure until $c_{11} > 0$ and $c_{1j} = 0$ for $j \geq 2$. As $c_{11} + \cdots + c_{1k}$ decreases by at least 1, the procedure terminates after a finite number of steps.

When $c_{11} > 0$ and $c_{1j} = 0$ for $j \geq 2$, we apply operation (iii) in Definition 5.2 to make $0 \leq b_{1j} < c_{11}$ for $j = 1, \ldots, d - k$ where d is the number of columns in A. Then the submatrix consisting of $\begin{bmatrix} D & 0 \end{bmatrix}$ and the row $\begin{bmatrix} b_{11} & \cdots & b_{1d-k} & c_{11} & 0 & \cdots & 0 \end{bmatrix}$ is in Hermite normal form. □

Remark 5.4 ([KB79])**.** The number of unimodular column operations needed in the proof is polynomially bounded.

2.2 Unimodular matrices

We say that an $m \times d$ matrix U is **unimodular** if

- it has full row rank m,
- all its entries are integers,
- every $m \times m$ submatrix B of U satisfies $\det(B) \in \{-1, 0, 1\}$.

Hence, a square matrix U is unimodular if all its entries are integers and $\det(U) \in \{-1, 1\}$ (Note: a square unimodular matrix is non-singular). In fact, the unimodular column operations can be performed by unimodular matrices.

First, interchanging column i and column j in matrix A can be done by computing AU where matrix U is given by

- $u_{kk} = 1$ for $k \in \{1, \ldots, d\} \setminus \{i, j\}$,
- $u_{ij} = u_{ji} = 1$,
- the other entries are 0.

Here, U is obtained after interchanging the ith column and the jth column of the $d \times d$ identity matrix. Then U is a $d \times d$ square matrix with $\det(U) = -1$, so U is unimodular.

Second, multiplying column i by -1 can be performed by AU where matrix U is obtained after replacing the ith diagonal entry of the $d \times d$ identity matrix by -1. Then $\det(U) = -1$, and thus U is unimodular.

The third operation of adding an integer multiple, say α, of column i to column j can be performed by AU where matrix U is obtained after replacing the entry of the $d \times d$ identity matrix at row i and column j by α. Note that all entries of U are integers, taking a value in $\{0, 1, \alpha\}$. Moreover, the diagonal entries of U are still all 1, which means $\det(U) = 1$. Therefore, U is unimodular.

The following lemma explains useful properties of unimodular matrices.

Lemma 5.5. *Let U be a $d \times d$ nonsingular matrix. Then the following statements are equivalent*

(i) U is unimodular,

(ii) both U and U^{-1} have all integral entries,

(iii) U^{-1} is unimodular,

(iv) Ux is integral if and only if x is integral for all $x \in \mathbb{R}^d$,

(v) U is obtained from the $d \times d$ identity matrix by a sequence of unimodular column operations.

2.3 Solving the equality constrained integer feasibility problem

Now, let's get back to solving (∗).

1. First, we check whether the constraint matrix A is of full row rank. If A is not of full row rank, then $Ax = b$ has some redundant constraints or $Ax = b$ induces an infeasible system.

For example,

$$\begin{aligned} 2x_1 + x_2 + 2x_3 + x_4 &= 5, \\ x_1 + 2x_2 + x_3 + 2x_4 &= 10, \\ x_1 + x_2 + x_3 + x_4 &= 5 \end{aligned}$$

corresponds to a constraint matrix that does not have full row rank. The system can be rewritten as

$$\begin{bmatrix} 2 & 1 & 2 & 1 \\ 1 & 2 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} 5 \\ 10 \\ 5 \end{bmatrix}.$$

Here, the third constraint is obtained from adding the first two constraints and dividing the resulting one by 3. In this case, we may remove the third constraint. Note that if the right-hand side of the third constraint were not 5, then the system would be infeasible because the first two constraints imply $x_1 + x_2 + x_3 + x_4 = 5$.

In case when A is not of full row rank and $Ax = b$ has some redundant constraints, we remove the redundant constraints so that the remaining constraint matrix is of full row rank. Thus we may assume that A is of full row rank.

2. The second step is to compute the Hermite normal form of A, given by $H = AU$ where U is some unimodular matrix.

Then $Ax = b$ is equivalent to

$$Ax = AUU^{-1}x = HU^{-1}x.$$

Here, as H is in Hermite normal form, H given by $H = \begin{bmatrix} D & 0 \end{bmatrix}$ for some nonsingular matrix D.

3. Let y be the subvector of $U^{-1}x$ corresponding to matrix D. Then $HU^{-1}x = b$ is equivalent to $Dy = b$, which gives $y = D^{-1}b$. Here,

$$U^{-1}x = \begin{bmatrix} D^{-1}b \\ z \end{bmatrix}$$

for any $z \in \mathbb{Z}^{d-m}$ corresponds to a candidate solution.

4. Compute

$$x = U \begin{bmatrix} D^{-1}b \\ z \end{bmatrix}$$

for any $z \in \mathbb{Z}^{d-m}$. Here, $x \in \mathbb{Z}^d$ if and only if $D^{-1}b \in \mathbb{Z}^m$.

References

[KB79] R. Kannan and A. Bachem. Polynomial algorithms for computing the Smith and Hermite normal forms of an integer matrix, 1979. 5.4