3

Dissecting matrices with Gauss-Jordan elimination

One ring to rule them all,
one ring to find them,
one ring to bring them all
and in the darkness bind them.

Sauron

3.1 An algorithm to rule them all

The aim of this chapter is to present an algorithm, known as *Gauss-Jordan elimination*, that takes a matrix A as input and returns: i) a basis of $C(A)$, ii) a basis of $N(A)$.

3.2 Elementary column operations

Given an $n \times m$ matrix A, we will denote its j-th column as A_j. We define the *elementary column operations* as either of the following ways to transform A:

- Rearrange (permute) the columns of A
- Replace A_j by $A_j + \lambda A_i$, with $i \neq j$ and $\lambda \in \mathbb{R}$
- Replace A_j by λA_j, with $\lambda \neq 0$

Interestingly, when we apply an elementary column operation on a matrix A, under the hood we are multiplying A by some matrix E on the right. Can you tell which matrices correspond to which operations?

Proposition 7. *If A' is obtained from A by applying an elementary column operation, then $C(A) = C(A')$.*

3.3 Reduced column echelon form

We need to introduce a couple of technical definitions that will help us be precise about the workings of our master algorithm.

We define the *leading index* of a column of A to be the smallest index with a non-zero entry, in case the column is not the zero vector, or ∞ if the column is zero. For a matrix A, we will denote the leading index of the j-th column as $\ell_j(A)$.

The *leading entry* of a non-zero column is simply the entry corresponding to the leading index.

3.3.1 Example

The matrix

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 2 & 5 & 0 \\ 3 & 6 & 0 \\ 4 & 7 & 8 \end{bmatrix}$$

has leading indices $\ell_1(A) = 1$, $\ell_2(A) = 2$ and $\ell_3(A) = 4$ with corresponding leading entries 1, 5 and 8, respectively.

An $n \times m$ matrix is said to be in *column echelon form* if its leading indexes have the following pattern:

$$\ell_1(A) < \ldots < \ell_k(A) = \ldots = \ell_m(A) = \infty,$$

for some $k \leq m$.

A matrix is said to be in *reduced column echelon form* if the following conditions hold:

- It is in column echelon form;
- The leading entries are 1;
- The leading entries are the only non-zero entries in their respective rows.

The matrix A in the previous example is in column echelon form, but not in reduced column echelon form.

3.3.2 Example

The matrix

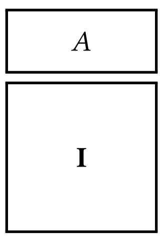
$$B = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

is in reduced column echelon form.

3.4 Gauss-Jordan Elimination

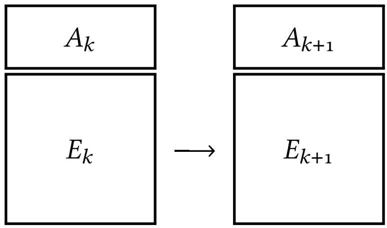
3.4.1 Set-up (initialization)

We will stack our input matrix A on top of an identity matrix with as many columns as A has:



3.4.2 Steps (loop)

At each step of the algorithm we will apply some elementary column operation to A and the bottom matrix simultaneously.



3.4.3 Goal (stop condition)

The goal is to reach a transform of A (top matrix) in column echelon form. Depending on the type of problem we want to addess it might be preferable to reach a reduced echelon form instead of a (vanilla) echelon form.

3.5 Interpreting the result

The output of the algorithm is rich in information about the matrix A.

Proposition 8. *Suppose that A is an $n \times m$ matrix, A^* is an echelon form obtained from A by repeatedly applying elementary column operations and E is the matrix obtained from $\boldsymbol{I}$ by applying the same operations. Then the following holds:*

- $A^* = AE$
- *The non-zero columns of A^* are a basis of $C(A)$.*
- *If A^* has zero columns, the corresponding columns of E are a basis of $N(A)$.*

- *$E = E_1 E_2 \dots E_s$ where E_i is the matrix corresponding to the elementary column operation used in step i.*

In view of all this, we can already state one of the most important results of the course.

Theorem 1 (Fundamental Theorem of Linear Algebra)**.** *For any $n \times m$ matrix A,*

$$m = \boldsymbol{rank}(A) + \boldsymbol{dim}(N(A))$$

3.6 Inverse of a matrix

When A is a square matrix, we may ask whether there is another matrix B of the same size, such that $AB = \mathbf{I}$. As a matter of fact, if such a matrix exists, it is the only matrix with this property and also $BA = \mathbf{I}$ must hold, too (why?). We call such matrix B the *inverse of A* and we typically denote it by A^{-1}.

We say that a square matrix A is *invertible* if it has an inverse.

Proposition 9. *If A is $n \times n$ invertible, then $\boldsymbol{rank}(A) = n$.*

Since $AA^{-1} = \mathbf{I}$ we can see that the columns of $\mathbf{I}$, which are a basis of $\mathbb{R}^n$, can be expressed as linear combinations of the columns of A, so the columns of A must also be a basis of $\mathbb{R}^n$, then $\mathbf{rank}(A) = n$.

The square matrices with this property are known as *full-rank.*

Interestingly, Gauss-Jordan elimination gives us a constructive way to check if a square matrix is A invertible and compute A^{-1} in one go.

Proposition 10 (Inverses via Gauss-Jordan Elimination)**.** *Suppose that we do Gauss-Jordan elimination on A until we get a **reduced echelon form** A^* and transformation matrix E. Then if A is full rank, then A^* is the identity matrix and $E = A^{-1}$.*

Then if A is a square, full-rank matrix, we can compute the inverse.

Proposition 11. *Given a square matrix A of size n, the following are equivalent:*

- *A is invertible*
- *A is full rank*
- *The columns of A are a basis of $\mathbb{R}^n$*
- *The rows of A are a basis of $\mathbb{R}^n$*

3.6.1 Remark

If two square matrices of the same size A, B are invertible, then their product is invertible, too:

$$(AB)^{-1} = B^{-1}A^{-1}$$

If a matrix A is invertible, then its transpose A^t is also invertible and, moreover

$$(A^t)^{-1} = (A^{-1})^t.$$

The elementary column operation matrices are all invertible. Consequently, the Gauss-Jordan transformation matrices E are also invertible, since they are products of elementary column operation matrices.

3.6.2 Remark

Multiplying some vector by the inverse of a matrix is equivalent to solve a system of linear equations. Whenever you are trying to solve a system of linear equations it should ring a bell that an inverse matrix multiplication is involved.

3.7 The CR factorization

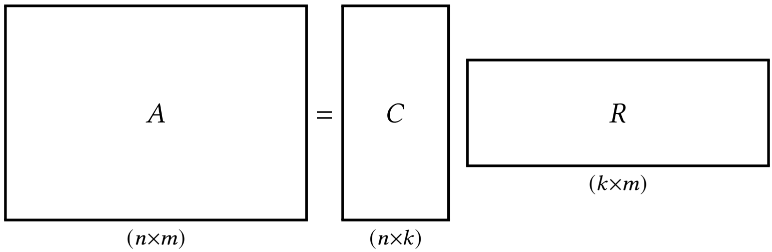
As we have just seen, for any $n \times m$ matrix A of rank k, we have that $A^* = AE$ where A^* is an $n \times m$ echelon form, possibly with the last $m - k$ columns zero, and E is an $m \times m$ invertible matrix.

Because E is invertible, we can write

$$A = A^* E^{-1}.$$

Now, the result of this product will be exactly the same if we get rid of the zero columns of A^* and we also get rid of the last $m - k$ rows of E^{-1}. If we denote C the pruned version of A^* and R the pruned version of E^{-1}, we have proved the following:

Proposition 12. *Any matrix A with rank k can be expressed* *as a product of matrices* $C \in \mathbb{R}^{n\times k}$ *and* $R \in \mathbb{R}^{k\times m}$*:*



In particular, the columns of C *are a basis of* $C(A)$ *and the rows of* R *are a basis of* $R(A)$*.*

3.8 The determinant

The *determinant* of a square matrix A, denoted $\mathbf{det}(A)$, is a number that can be defined recursively as follows.

- If A is a 1×1 matrix $[a_{11}]$, $\mathbf{det}(A) = a_{11}$
- If A is an $n \times n$ matrix, then for any $1 \leq j \leq n$

$$\mathbf{det}(A) = \sum_{i=1}^{n} a_{ij}(-1)^{i+j}\mathbf{det}(A_{ij})$$

  where A_{ij} denotes the $(n-1) \times (n-1)$ matrix derived from A where the i-th row and j-th column have been removed.

The determinant has many interesting properties and can be very handy to solve many linear algebra questions. For the time being, we will underscore its importance to assert linear independence.

Proposition 13. *Given a square matrix A, the following are equivalent:*

- $\boldsymbol{det}(A) \neq 0$
- *A is full rank*
- *A is invertible*

Here are a bunch of interesting properties of the determinant:

Proposition 14. *Let A and B be square matrices of the same size:*

- *The determinant is linear column by column (and row by row).*
- $\boldsymbol{det}(AB) = \boldsymbol{det}(A)\boldsymbol{det}(B)$
- $\boldsymbol{det}(A^t) = \boldsymbol{det}(A)$
- *If A is invertible, $\boldsymbol{det}(A^{-1}) = 1/\boldsymbol{det}(A)$*
- *If A^σ is the matrix obtained by rearranging the columns (or rows) of A with a permutation σ, then*

  $$\boldsymbol{det}(A^\sigma) = sgn(\sigma)\boldsymbol{det}(A),$$

  where sgn(σ) is the sign of the permutation σ.

That the determinant is linear column by column means that:

- It preserves sums column by column:

  $$\mathbf{det}([\ldots \mid a+b \mid \ldots]) =$$

$$= \mathbf{det}([\ldots \mid a \mid \ldots]) + \mathbf{det}([\ldots \mid b \mid \ldots])$$

- It preserves scaling column by column:

$$\mathbf{det}([\ldots \mid \lambda a \mid \ldots]) = \lambda\, \mathbf{det}([\ldots \mid a \mid \ldots])$$

3.9 TLDR;

We developed a technique that is very useful to study any matrix, known as Gauss-Jordan elimination. In its essence, the method transforms step by step our input matrix until we get to another matrix from which the dependence structure of the input matrix is much easier to read: at each step we are only allowed to use transformations that do not change the column space. With this method we can easily get a complete characterization of the column space and the null space. In addition, this method gives us, free of charge, a method to compute the inverse of input invertible square matrices. Square matrices are an important particular case, where we can also resort to other, more sophisticated techniques, like the determinant, that will lets us automatically read whether the input matrix is full rank or not.