

# Gaussian Elimination

Siddhartha Sahi  
Math Department, Rutgers University

February 14, 2007

## 1 The algorithm

**Gaussian Elimination** (GE) is an procedure for converting a matrix into its **row echelon form** (REF) and **reduced row echelon form** (RREF), using row operations. The algorithm has three parts:

1. The forward pass
2. Normalization
3. The backward pass

### 1.1 Forward pass

During the forward pass we use row operations on successively smaller matrices. The procedure is as follows:

1. Determine the **first nonzero column**, called the **pivot column**.
2. If necessary, do a row swap (denoted  $R_i \leftrightarrow R_j$ ) to bring a nonzero entry to the top of the pivot column, called the **pivot position**
3. Subtract multiples of the top row called the **pivot row**, from the lower rows (denoted  $R_i - aR_j$ ) to create zeros below the pivot position.
4. Mentally “cross out” the pivot row to get a smaller matrix.
5. Repeat steps 1, 2, 3, 4 for the smaller matrix until done.

Here is an example of the forward pass applied to the matrix

$$\begin{bmatrix} 1 & 4 & 2 & 2 & 2 \\ 2 & 8 & 4 & 1 & 4 \\ 2 & 7 & 4 & 2 & 2 \end{bmatrix}$$

The pivot column is column 1. Step 2 is unnecessary. In Step 3 we get

$$\begin{bmatrix} \mathbf{1} & 4 & 2 & 2 & 2 \\ \mathbf{2} & 8 & 4 & 1 & 4 \\ \mathbf{2} & 7 & 4 & 2 & 2 \end{bmatrix} \begin{array}{l} R_2 - 2R_1 \\ R_3 - 2R_1 \end{array} \rightarrow \begin{bmatrix} \mathbf{1} & 4 & 2 & 2 & 2 \\ \mathbf{0} & 0 & 0 & -3 & 0 \\ \mathbf{0} & -1 & 0 & -2 & -2 \end{bmatrix}$$

We mentally cross out the first row and work with the smaller matrix (last two rows). The pivot column is column 2. In step 2 we swap rows 2 and 3. Step 3 is unnecessary. We get

$$\begin{bmatrix} \mathbf{1} & \mathbf{4} & 2 & 2 & 2 \\ \mathbf{0} & \mathbf{0} & 0 & -3 & 0 \\ \mathbf{0} & \mathbf{-1} & 0 & -2 & -2 \end{bmatrix} \begin{array}{l} R_2 \leftrightarrow R_3 \end{array} \rightarrow \begin{bmatrix} \mathbf{1} & \mathbf{4} & 2 & 2 & 2 \\ \mathbf{0} & \mathbf{-1} & 0 & -2 & -2 \\ \mathbf{0} & \mathbf{0} & 0 & -3 & 0 \end{bmatrix}$$

We mentally cross out the second row also and work with the still smaller matrix (only last row). The pivot column is column 4, and no steps are needed. So we get

$$\begin{bmatrix} \mathbf{1} & \mathbf{4} & 2 & \mathbf{2} & 2 \\ \mathbf{0} & \mathbf{-1} & 0 & \mathbf{-2} & -2 \\ \mathbf{0} & \mathbf{0} & 0 & \mathbf{-3} & 0 \end{bmatrix}$$

where all the pivot columns have been highlighted for convenience.

## 1.2 Normalization

After the forward pass has been completed, we multiply/divide the pivot rows by an appropriate scalar (denoted  $aR_i$ ) to make all pivot entries equal to 1. In the example above this gives

$$\begin{bmatrix} \mathbf{1} & \mathbf{4} & 2 & \mathbf{2} & 2 \\ \mathbf{0} & \mathbf{-1} & 0 & \mathbf{-2} & -2 \\ \mathbf{0} & \mathbf{0} & 0 & \mathbf{-3} & 0 \end{bmatrix} \begin{array}{l} -R_2 \\ -1/3R_3 \end{array} \rightarrow \begin{bmatrix} \mathbf{1} & \mathbf{4} & 2 & \mathbf{2} & 2 \\ \mathbf{0} & \mathbf{1} & 0 & \mathbf{2} & 2 \\ \mathbf{0} & \mathbf{0} & 0 & \mathbf{1} & 0 \end{bmatrix}$$

## 1.3 Backward pass

In this stage we use row operations to create zeroes above the pivots. The procedure is as follows:

1. Start with the last pivot and subtract multiples of the pivot row from upper rows ( $R_i - aR_j$ ) to create zeroes above the pivot.
2. Go successively to the previous pivots and repeat the procedure.

For the example we get

$$\begin{bmatrix} \mathbf{1} & \mathbf{4} & 2 & \mathbf{2} & 2 \\ \mathbf{0} & \mathbf{1} & 0 & \mathbf{2} & 2 \\ \mathbf{0} & \mathbf{0} & 0 & \mathbf{1} & 0 \end{bmatrix} \begin{array}{l} R_2 - 2R_3 \\ R_1 - 2R_3 \end{array} \rightarrow \begin{bmatrix} \mathbf{1} & \mathbf{4} & 2 & \mathbf{0} & 2 \\ \mathbf{0} & \mathbf{1} & 0 & \mathbf{0} & 2 \\ \mathbf{0} & \mathbf{0} & 0 & \mathbf{1} & 0 \end{bmatrix} \begin{array}{l} R_1 - 4R_2 \end{array} \rightarrow \begin{bmatrix} \mathbf{1} & \mathbf{0} & 2 & \mathbf{0} & -6 \\ \mathbf{0} & \mathbf{1} & 0 & \mathbf{0} & 2 \\ \mathbf{0} & \mathbf{0} & 0 & \mathbf{1} & 0 \end{bmatrix}$$

## 1.4 Another example

We give another example of GE, this time simply indicating the steps

$$\begin{aligned} & \begin{bmatrix} \mathbf{1} & 4 & 2 & 2 & 2 \\ \mathbf{2} & 8 & 3 & 5 & 1 \end{bmatrix} \xrightarrow{R_2 - 2R_1} \begin{bmatrix} \mathbf{1} & 4 & \mathbf{2} & 2 & 2 \\ \mathbf{0} & 0 & -\mathbf{1} & 1 & -3 \end{bmatrix} \\ & \xrightarrow{-R_2} \begin{bmatrix} \mathbf{1} & 4 & \mathbf{2} & 2 & 2 \\ \mathbf{0} & 0 & \mathbf{1} & -1 & 3 \end{bmatrix} \xrightarrow{R_1 - 2R_2} \begin{bmatrix} \mathbf{1} & 4 & \mathbf{0} & 4 & -4 \\ \mathbf{0} & 0 & \mathbf{1} & -1 & 3 \end{bmatrix} \end{aligned}$$

## 1.5 Echelon forms

The matrix obtained at the end of the forward pass is called the row echelon form (REF) of  $A$ . It has the following properties:

1. The leading non-zero entries in each row are staggered to the down and right, as in a staircase (echelon).
2. Any entirely zero rows are at the bottom of the matrix.

The matrix obtained at the end of the backward pass is called the reduced row echelon form (RREF) of  $A$ . It has two further properties:

3. The leading non-zero entries are 1,
4. All other entries in the column of a leading entry are 0.

## 1.6 Rank and nullity

The number of pivots in RREF of  $A$  is called the **rank** of the matrix  $A$ .

The number of non-pivot columns of  $A$  is called the **nullity** of  $A$ .

Note that

$$\text{rank of } A + \text{nullity of } A = \text{number of columns of } A$$

## 2 Solving systems of equations

### 2.1 Augmented matrix

Consider a system of linear equations of the form

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n &= b_2 \\ &\vdots \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n &= b_m \end{aligned}$$

where  $a_{ij}$ ,  $b_i$  are some constants and  $x_i$  are unknown variables.

This can be rewritten in matrix-vector form  $Ax = b$  where

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}, \quad x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}, \quad b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

To solve such a system we apply GE to its **augmented matrix**.

$$[A \ b] = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} & b_1 \\ a_{21} & a_{22} & \cdots & a_{2n} & b_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} & b_m \end{bmatrix}$$

### 2.2 Solution using GE

Note that the first  $n$  columns of the augmented matrix correspond to the variables, and the rows correspond to the equations. The various row operations of GE correspond to swapping, combining and rescaling equations. These operations do not change the solutions of the system.

The solutions can be read off from the RREF matrix of  $[Ab]$  as follows:

1. The pivot column variables are called **basic**, the others are called **free**.
2. If  $x_i, x_j, \dots, x_k$  are the free variables, then the general solution is

$$-x_i \mathbf{u}_i - x_j \mathbf{u}_j - \cdots - x_k \mathbf{u}_k + \mathbf{u}_{n+1}$$

3. where  $\mathbf{u}_{n+1}, \mathbf{u}_i, \mathbf{u}_j, \dots, \mathbf{u}_k \in \mathbb{R}^n$  are constructed from RREF in two steps
  - For each  $i \leq n$ , the  $i$ th coordinate of each  $\mathbf{u}_i$  is set to  $-1$ , all the other “free” coordinates of the  $\mathbf{u}$ 's are set to 0.
  - The “basic” coordinates of  $\mathbf{u}$ 's are obtained by copying the corresponding column from the RREF into the basic slots, in order.

**Example 1** Consider the system

$$\begin{aligned}x_1 + 4x_2 + 2x_3 + 2x_4 &= 2 \\2x_1 + 8x_2 + 4x_3 + x_4 &= 4 \\2x_1 + 7x_2 + 4x_3 + 2x_4 &= 2\end{aligned}$$

We have already computed the RREF of the augmented matrix:

$$\begin{bmatrix} \mathbf{1} & \mathbf{0} & 2 & \mathbf{0} & -6 \\ \mathbf{0} & \mathbf{1} & 0 & \mathbf{0} & 2 \\ \mathbf{0} & \mathbf{0} & 0 & \mathbf{1} & 0 \end{bmatrix}$$

There is only one free variable  $x_3$ , so the solution is of the form

$$-x_3 \mathbf{u}_3 + \mathbf{u}_5$$

We fill in the free components (3rd component) of  $\mathbf{u}_3$  and  $\mathbf{u}_5$  as follows:

$$\mathbf{u}_3 = \begin{bmatrix} \\ -1 \\ \end{bmatrix}, \mathbf{u}_5 = \begin{bmatrix} \\ 0 \\ \end{bmatrix}$$

Then we copy the basic components (1st, 2nd, 4th) from the 3rd and 5th columns as follows:

$$\mathbf{u}_3 = \begin{bmatrix} 2 \\ 0 \\ 0 \\ \end{bmatrix}, \mathbf{u}_5 = \begin{bmatrix} -6 \\ 2 \\ 0 \\ \end{bmatrix}$$

So the final answer is

$$-x_3 \begin{bmatrix} 2 \\ 0 \\ -1 \\ 0 \\ \end{bmatrix} + \begin{bmatrix} -6 \\ 2 \\ 0 \\ 0 \\ \end{bmatrix} = x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \\ \end{bmatrix} + \begin{bmatrix} -6 \\ 2 \\ 0 \\ 0 \\ \end{bmatrix}$$

Why does this work? The reason is as follows: The RREF system can be written as

$$\begin{aligned}x_1 + 2x_3 &= -6 \\x_2 &= 2 \\x_4 &= 0\end{aligned}$$

Thus we get

$$\begin{bmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{bmatrix} = \begin{bmatrix} -2x_3 - 6 \\ 2 \\ x_3 \\ 0 \end{bmatrix} = x_3 \begin{bmatrix} -2 \\ 0 \\ 1 \\ 0 \end{bmatrix} + \begin{bmatrix} -6 \\ 2 \\ 0 \\ 0 \end{bmatrix}$$

**Example 2** For another example consider the system

$$\begin{aligned}x_1 + 4x_2 + 2x_3 + 2x_4 &= 2 \\2x_1 + 8x_2 + 3x_3 + 5x_4 &= 1\end{aligned}$$

Here too we have calculated the RREF of the augmented matrix:

$$\begin{bmatrix} \mathbf{1} & 4 & \mathbf{0} & 4 & -4 \\ \mathbf{0} & 0 & \mathbf{1} & -1 & 3 \end{bmatrix}$$

The free variables are  $x_2$  and  $x_4$ , and the solution is

$$\begin{aligned}& -x_2 \left( \begin{bmatrix} -1 \\ 0 \end{bmatrix} + \begin{bmatrix} 4 \\ 0 \end{bmatrix} \right) - x_4 \left( \begin{bmatrix} 0 \\ -1 \end{bmatrix} + \begin{bmatrix} 4 \\ -1 \end{bmatrix} \right) + \left( \begin{bmatrix} 0 \\ 0 \end{bmatrix} + \begin{bmatrix} -4 \\ 3 \end{bmatrix} \right) \\&= -x_2 \begin{bmatrix} 4 \\ -1 \\ 0 \\ 0 \end{bmatrix} - x_4 \begin{bmatrix} 4 \\ 0 \\ -1 \\ -1 \end{bmatrix} + \begin{bmatrix} -4 \\ 0 \\ 3 \\ 0 \end{bmatrix} = x_2 \begin{bmatrix} -4 \\ 1 \\ 0 \\ 0 \end{bmatrix} + x_4 \begin{bmatrix} -4 \\ 0 \\ 1 \\ 1 \end{bmatrix} + \begin{bmatrix} -4 \\ 0 \\ 3 \\ 0 \end{bmatrix}\end{aligned}$$

### 2.3 Calculation check

The vector  $u_{n+1}$  is always a solution of the original system (it corresponds to setting all the free variables equal to 0 in the general solution). However the vectors  $u_i, u_j$  etc. are not solutions of the system, but instead satisfy the *homogeneous* system

$$A\mathbf{x} = 0$$

Therefore we can “check” our calculations by verifying that

$$A\mathbf{u}_{n+1} = b \text{ and } A\mathbf{u}_i = A\mathbf{u}_j = \dots = 0$$

For the second example we get

$$\begin{aligned}(-4) + 4(0) + 2(3) + 2(0) &= 2 \text{ and } 2(-4) + 8(0) + 3(3) + 5(0) = 1 \\(-4) + 4(1) + 2(0) + 2(0) &= 0 \text{ and } 2(-4) + 8(0) + 3(3) + 5(0) = 0 \\(-4) + 4(0) + 2(1) + 2(1) &= 0 \text{ and } 2(-4) + 8(0) + 3(1) + 5(1) = 0\end{aligned}$$

## 2.4 Consistency and number of solutions

When does a system not have any solutions?

This happens when GE produces a pivot in the last column of the augmented matrix. This means that we get a row of the form

$$0 \ 0 \ \dots \ 0 \ 1$$

which corresponds to the “impossible” equation

$$0 = 1.$$

Since the rank of a matrix is the number of pivots in its RREF. Therefore the following are equivalent

1. The system  $Ax = b$  is consistent (has solutions)
2. There is no pivot in the last column of the RREF of  $[Ab]$
3.  $\text{rank}(A) = \text{rank}([Ab])$

In general there are three possibilities for the number of solutions:

1. There are infinitely many solutions, depending on some free variables.
2. There is exactly one solution and no free variables.
3. There are no solutions – we say the system is *inconsistent*.