1 Recap

- Approximation problem and examples.
- Left and right inverses, inverses of onto linear maps.

2 Linear equations

Recall that linear systems of equations like $2x+3y+z=20, x+y-z=\pi$ can be written using matrices as $AX=b$ where $A=\begin{bmatrix} 2 & 3 & 1 \\ 1 & 1 & -1 \end{bmatrix}$, $X=\begin{bmatrix} x \\ y \\ z \end{bmatrix}$, and $b=\begin{bmatrix} 20 \\ \pi \end{bmatrix}$.

More generally, $\sum_j A_{ij}x_j = b_i$, i.e., $AX=b$ represents a system of linear equations. The matrix A is called the *coefficient matrix*. As mentioned earlier, systems can fail to have solutions or even have infinitely many solutions.

If $b=0$, then the system $AX=0$ is called a *homogeneous* system. Recall that if $AX_0=b$, then *any* other solution to $AX=b$ is of the form $X=X_0+N$ where $AN=0$. So it suffices to solve $AN=0$ and find a *single* solution to $AX=b$.

So how does one solve linear equations? One is allowed to

- Interchange equations.
- Multiply both sides of *an* equation by a *nonzero* scalar.
- Add one equation to a multiple of another.

The high-school idea is to eliminate a few variables and solve for the rest by "back-substitution". This idea was formalised and used to great effect by Gauss and Jordan.

3 Gauss-Jordan elimination

Firstly, in the example above the variables x, y, z are distractions. After all, we only care about manipulating the *coefficients*.

So we define the *augmented matrix* $[A|b]$ by simply adding b as a column to A. Notice that the three "legal" operations alluded to above are:

- Interchanging the rows of $[A|b]$. (Each row corresponds to an equation.)
- Multiply any row by a nonzero scalar.
- Add a row to a multiple of another.

These operations are called *elementary row operations*. The aim is to do these operations and bring the matrix to a special form (known as the *row-echelon* form).

A matrix C is said to be in the row-echelon form if *below* the *first* non-zero entry of every row all the elements are *zero*. The point is to solve the *last* non-trivial equation and back-substitute to solve the rest.

Examples and non-examples of row-echelon matrices

- $\begin{bmatrix} 2 & 3 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$ is *not* in the row-echelon form.

- $\begin{bmatrix} \pi & 2 \\ 0 & e \end{bmatrix}$ is in the row-echelon form.

- $\begin{bmatrix} \sqrt{-1} & 1 \\ 0 & 1 \\ 0 & 0 \end{bmatrix}$ is in the row-echelon form.

An $m \times n$ matrix A is said to be in the reduced row-echelon form if it is in the row-echelon form, each pivot is 1, and the column containing each pivot has only zeroes in the other entries.
If A is in the row-echelon form then it can be reduced to the reduced row-echelon form easily using further row operations.

A theorem of Gauss and Jordan is: Every $m \times n$ matrix A with entries in a field $\mathbb{F}$ can be row-reduced to a *unique* reduced row-echelon form.
The theorem can be proven using induction on the number of rows. Two crucial observations are:

- Elementary row operations can be reversed, i.e., run backwards.

- If one gets B from A using elementary row operations, then each row of B is a linear combination of rows of A. (The linear span of rows of a matrix A is called the *row space* of A. Likewise, that of the columns is called the *column space*.)

Row-reduction does not change the row space (HW). We shall not prove the theorem. Instead we shall illustrate its application to linear equations using examples.

4 The row-reduction algorithm

- Identify the left-most pivot among all rows. Suppose it occurs in the i^{th} row.

- Interchanging rows, make sure that R_i is the first row.

- Divide out the first-row pivot to make it 1.

- “Clear” everything below the first-row pivot using row operations.

- By induction/recursion/“Rinse and repeat” the $(m-1) \times n$ matrix of the next $m-1$ rows can be assumed to be in the required form.

- Clear the elements in the first row using the pivots in the other rows. (On a computer, you can implement it iteratively or recursively.)

To solve $Ax = b$, consider the augmented matrix $[A|b]$, and row-reduce it to its RREF $[\tilde{A}|\tilde{b}]$.
If any row of $\tilde{A}$ is 0, but the corresponding entry of b is not, then the system is inconsistent.
If it is consistent, starting from the bottom of $\tilde{A}$ solve for the first non-zero pivoted variable.
Inductively/recursively, solve for the other pivoted variables.

5 Examples of solving equations

- Solve: $2x - 5y + 4z = -3, x - 2y + z = 5, x - 4y + 6z = 10$.
  The augmented matrix is $\left[\begin{array}{ccc|c} 2 & -5 & 4 & -3 \\ 1 & -2 & 1 & 5 \\ 1 & -4 & 6 & 10 \end{array}\right]$
  $R_1 \to R_1/2$ gives $\left[\begin{array}{ccc|c} 1 & -\frac{5}{2} & 2 & -\frac{3}{2} \\ 1 & -2 & 1 & 5 \\ 1 & -4 & 6 & 10 \end{array}\right]$.
  Now we "clear" the first column through $R_2 \to R_2 - R_1,\ R_3 \to R_3 - R_1$ to get $\left[\begin{array}{ccc|c} 1 & -\frac{5}{2} & 2 & -\frac{3}{2} \\ 0 & \frac{1}{2} & -1 & \frac{13}{2} \\ 0 & -\frac{3}{2} & 4 & \frac{23}{2} \end{array}\right]$.

  Rinse and repeat: $R_2 \to 2R_2$ and then $R_3 \to R_3 + \frac{3}{2}R_2,\ R_1 \to R_1 + \frac{5}{2}R_2$ give $\left[\begin{array}{ccc|c} 1 & 0 & -3 & 31 \\ 0 & 1 & -2 & 13 \\ 0 & 0 & 1 & 31 \end{array}\right]$.
  It is not in RREF but, we can solve now itself: $z = 31$, $y = 13 + 2z = 75$, and $x = 3z + 31 = 124$.

- Solve: $x - 2y + z - u + v = 5, 2x - 5y + 4z + u - v = -3, x - 4y + 6z - v + 2u = 10$.
  The augmented matrix is $\left[\begin{array}{ccccc|c} 1 & -2 & 1 & -1 & 1 & 5 \\ 2 & -5 & 4 & 1 & -1 & -3 \\ 1 & -4 & 6 & 2 & -1 & 10 \end{array}\right]$.
  We clear the first column through $R_2 \to R_2 - 2R_1$, $R_3 \to R_3 - R_1$ to get $\left[\begin{array}{ccccc|c} 1 & -2 & 1 & -1 & 1 & 5 \\ 0 & -1 & 2 & 3 & -3 & -13 \\ 0 & -2 & 5 & 3 & -2 & 5 \end{array}\right]$

  Now we normalise the second row: $R_2 \to -R_2$ and then clear the second column: $R_3 \to R_3 + 2R_2,\ R_1 \to R_1 + 2R_2$ to get $\left[\begin{array}{ccccc|c} 1 & 0 & -3 & -7 & 7 & 31 \\ 0 & 1 & -2 & -3 & 3 & 13 \\ 0 & 0 & 1 & -3 & 4 & 31 \end{array}\right]$.
  We clear the third column: $R_2 \to R_2 + 2R_3$, $R_1 \to R_1 + 3R_3$ to get $\left[\begin{array}{ccccc|c} 1 & 0 & 0 & -16 & 19 & 124 \\ 0 & 1 & 0 & -9 & 11 & 75 \\ 0 & 0 & 1 & -3 & 4 & 31 \end{array}\right]$.
  Thus $z = 3u - 4v + 31$, $y = 9u - 11v + 75$, $x = 16u - 19v + 124$. That is, $(x, y, z, u, v) = (124, 75, 31, 0, 0) + u(16, 9, 3, 1, 0) + v(-19, -11, -4, 0, 1)$. $(124, 75, 31, 0, 0)$ is a *particular* solution, and when the matrix A is considered as a linear map, $(16, 9, 3, 1, 0)$ and $(-19, -11, -4, 0, 1)$ span the kernel .