MATH 471

Review of Linear Systems and Matrix Algebra

Our objective is to be able to efficiently solve large linear systems of equations on a computer. As we have seen, linear systems can crop up in approximating solutions to non-linear systems. Other uses we will have for these technique will arise when we discuss function approximation and approximating solutions to systems of ordinary differential equations.

In general we will be concerned with the solution to a systems of n equations and n unknowns

$$\begin{aligned}
a_{11}x_1 \;+\; a_{12}x_2 \;+\; a_{13}x_3 \;+\; \ldots + a_{1n}x_n &= b_1\\
a_{21}x_1 \;+\; a_{22}x_2 \;+\; a_{23}x_3 \;+\; \ldots + a_{2n}x_n &= b_2\\
a_{31}x_1 \;+\; a_{32}x_2 \;+\; a_{33}x_3 \;+\; \ldots + a_{3n}x_n &= b_3\\
\vdots \qquad &= \vdots\\
a_{n1}x_1 \;+\; a_{n2}x_2 \;+\; a_{n3}x_3 \;+\; \ldots + a_{nn}x_n &= b_n
\end{aligned}$$

where a_{ij} and b_i are constants and $x_1, x_2, x_3, \ldots, x_n$ are the unknowns we seek. There are three possibility; the system has no solution, the system has a unique solution, or the systems has infinitely many solutions. Matrix algebra will paly an essential role in investigating of this problem. The tools from matrix algebra will answer questions about uniqueness of solutions and will aid us in the development algorithms for determining solutions to linear systems of equations.

Review of Matrix Algebra

Before jumping into methods for solving systems of equations, it will be helpful to review some basics of matrix algebra, such as addition, multiplication, what is a matrix inverses, etc. Lets start with some examples of matrices.

A matrix is an ordered block of numbers consisting of rows and columns. Examples of three different matrices are

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 7 & 19 & \frac{1}{4} \\ 0 & 1 & 4 \end{bmatrix} \quad B = \begin{bmatrix} 4 & 5 & 0 \\ 1 & 2 & 7 \end{bmatrix} \quad C = \begin{bmatrix} 3 \\ \frac{1}{2} \\ 0 \\ 1 \end{bmatrix}.$$

Def. Row of a matrix: A row of a matrix is a line of number form left to right. An example of a row of a matrix is the bold text in matrix below,

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 7 & 19 & \frac{1}{4} \\ \mathbf{0} & \mathbf{1} & \mathbf{4} \end{bmatrix}.$$

Rows are label by indexes which start counting from 1. The bold text in our example is the third row of matrix A.

Def. Column of a matrix: A column of a matrix is a line of number form top to bottom. An example of a column of a matrix is the bold text in matrix below,

$$A = \begin{bmatrix} 1 & \mathbf{2} & 3 \\ 7 & \mathbf{19} & \frac{1}{4} \\ 0 & \mathbf{1} & 4 \end{bmatrix} .$$

Like rows, columns are label by indexes which start counting from 1. The bold text in our example is the second column of matrix A.

Def. Dimension of a matrix: The dimension of a matrix is defined to be the number of rows by the number of columns. For our three examples, A, B and C, the dimension are;

- Matrix A has 3 rows and 3 columns, so matrix A has dimension 3×3
- Matrix B has 2 rows and 3 columns, so matrix B has dimension 2×3
- Matrix C has 4 rows and 1 columns, so matrix C has dimension 4×1 .

Def. Transpose of a matrix: The transpose of a matrix A is a matrix which has interchanges the rows and columns of A. The transpose of a matrix is denoted by A^T. The transpose of the matrix B is

$$B^T = \begin{bmatrix} 4 & 5 & 0 \\ 1 & 2 & 7 \end{bmatrix}^T = \begin{bmatrix} 4 & 1 \\ 5 & 2 \\ 0 & 7 \end{bmatrix} ,$$

where row 1 of B has become column 1 of B^T and row 2 of B has become column 2 of B^T.

Observation: An important fact that can easily be shown is the transpose of the transpose of a matrix is the original matrix, i.e.,

$$\left(A^T\right)^T = A \ .$$

Def. A symmetric matrix: A matrix is said to be symmetric if $A = A^T$. Example, let A be

$$A = \begin{bmatrix} \mathbf{4} & \mathbf{5} & \mathbf{0} \\ \mathit{5} & \mathit{2} & \mathit{7} \\ 0 & 7 & 3 \end{bmatrix}$$

then A^T is

$$A^T = \begin{bmatrix} \mathbf{4} & \mathit{5} & 0 \\ \mathbf{5} & \mathit{2} & 7 \\ \mathbf{0} & \mathit{7} & 3 \end{bmatrix}$$

therefor $A = A^T$ and so A is a symmetric matrix.

Addition of two matrices

Matrix addition is defined to be termwise addition, i.e., given two matrices of the same dimension

$$A = \begin{bmatrix} a_{11} & a_{12} & \cdots \\ a_{21} & a_{22} & \cdots \\ \vdots & & \ddots \end{bmatrix} \; and \; B = \begin{bmatrix} b_{11} & b_{12} & \cdots \\ b_{21} & b_{22} & \cdots \\ \vdots & & \ddots \end{bmatrix}$$

we define their addition to be

$$A + B = \begin{bmatrix} a_{11} + b_{11} & a_{12} + b_{12} & \cdots \\ a_{21} + b_{21} & a_{22} + b_{22} & \cdots \\ \vdots & & \ddots \end{bmatrix} .$$

Lets consider a 2×2 example,

$$\begin{bmatrix} 2 & 1 \\ 3 & 4 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 7 & -4 \end{bmatrix} = \begin{bmatrix} 2+0 & 1+1 \\ 3+7 & 4+(-4) \end{bmatrix} = \begin{bmatrix} 2 & 2 \\ 10 & 0 \end{bmatrix} .$$

Multiplication of two matrices

Let$A$ and B be matrices. If A is a $m \times p$ matrix and B is a $p \times n$ matrix, then we define the product of A and B to be a matrix whose elements are given by

$$(AB)_{ij} = \sum_{k=1}^{p} a_{ik} b_{kj}, \; (1 \le i \le m) \, and \, (1 \le j \le n) \; .$$

where the index ij denotes the location of the element in terms of rows and columns. i designates the row and j designates the column. Lets consider two examples. Let the matrices A, B and C be given as

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix}_{2\times 3} \quad B = \begin{bmatrix} 2 & 3 \\ 1 & 1 \\ 2 & 4 \end{bmatrix}_{3\times 2} , C = \begin{bmatrix} 1 \\ \frac{1}{2} \\ 3 \end{bmatrix}_{3\times 1} ,$$

where I have indicated the dimension of the matrix by placing an appropriate subscript on each of the three matrices. Lets start by computing AB. Element $(AB)_{11}$ is then given by

$$(AB)_{11} = \sum_{k=1}^{3} a_{1k} b_{k1} = (1\cdot 2 + 2\cdot 1 + 3\cdot 2) \ .$$

Likewise, elements $(AB)_{12}$, $(AB)_{21}$,$(AB)_{22}$ are given by

$$(AB)_{12} = \sum_{k=1}^{3} a_{1k} b_{k2} = (1\cdot 3 + 2\cdot 1 + 3\cdot 4)$$

$$(AB)_{21} = \sum_{k=1}^{3} a_{2k} b_{k1} = (3\cdot 2 + 1\cdot 1 + 4\cdot 2)$$

$$(AB)_{22} = \sum_{k=1}^{3} a_{2k} b_{k2} = (3\cdot 3 + 1\cdot 1 + 4\cdot 4)$$

or the matrix AB is given by

$$AB = \begin{bmatrix} (AB)_{11} & (AB)_{12} \\ (AB)_{21} & (AB)_{22} \end{bmatrix}_{2\times 2} = \begin{bmatrix} 10 & 17 \\ 13 & 23 \end{bmatrix}_{2\times 2} .$$

The matrix AC is given by

$$AC = \begin{bmatrix} 1 & 2 & 3 \\ 2 & 1 & 4 \end{bmatrix}_{2\times 3} \begin{bmatrix} 1 \\ \frac{1}{2} \\ 3 \end{bmatrix}_{3\times 1} = \begin{bmatrix} 1\cdot 1 + 2\cdot \frac{1}{2} + 3\cdot 3 \\ 2\cdot 1 + 1\cdot \frac{1}{2} + 4\cdot 3 \end{bmatrix}_{2\times 1} = \begin{bmatrix} 11 \\ 15 \end{bmatrix}_{2\times 1}$$

Scalar Multiplication of a Matrix and a Constant

Let $c \in \mathbf{R}$ and let A be matrices an $n \times m$ matrix. Then we define a scalar matrix product to be the termwise product of the scalar with each element of the matrix A.

$$cA = c \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1m} \\ a_{21} & a_{22} & \dots & a_{2m} \\ \vdots & & \ddots & \\ a_{n1} & a_{n2} & \dots & a_{nm} \end{bmatrix} = \begin{bmatrix} ca_{11} & ca_{12} & \dots & ca_{1m} \\ ca_{21} & ca_{22} & \dots & ca_{2m} \\ \vdots & & \ddots & \\ ca_{n1} & ca_{n2} & \dots & ca_{nm} \end{bmatrix} .$$

Def. Identity matrix: The identity matrix is a square matrix with ones down the diagonal and zeros everywhere else.

$$I = \begin{bmatrix} 1 & 0 & \dots & 0 \\ 0 & 1 & \dots & 0 \\ \vdots & & \ddots & \\ 0 & 0 & \dots & 1 \end{bmatrix} .$$

Def. Matrix Inverse: Let A be a $n \times p$ matrix, let B be a $p \times n$ matrix and let C be a $p \times n$ matrix . B is said to be the left inverse of A if

$$B_{pn} A_{np} = I_{pp} \ .$$

Likewise, C is said to as the right inverse of A if

$$A_{np} C_{pn} = I_{nn} \ .$$

We will denote an inverse of A by A^{-1} and in particular we will denote the left and right inverse by A_L^{-1} and A_R^{-1} respectively. Example,

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ a & b \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} .$$

So we see that the first matrix is a left inverse for any matrix of the from of the second matrix.

THM. Square Matrix Inverse: A square matrix A has at most one right inverse

$$AA_R^{-1} = I \ .$$

THM. The Commutative Property of Square Matrix Inverses: If A and B are $n \times n$ matrices such that $AB = I$, then $BA = I$.
Proof: Given $AB = I$, to show that $BA = I$.
Let

$$C = BA - I + B \ .$$

Multiplying through by A gives

$$AC = ABA - AI + AB$$

and recalling that $AB = I$ gives

$$AC = IA - AI + I$$

or

$$AC = A - A + I = I \ .$$

Hence C must be a right inverse of A (by the uniqueness of square inverses). Therefor $C = B$ and hence equation

$$(C = BA - I + B) \ \Rightarrow \ (0 = BA - I) \ \Rightarrow (BA = I)$$

as was to be shown.

Def. Determent of a Matrix: For a 2×2 matrix

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

we define the determent to be

$$|A| = \begin{vmatrix} a & b \\ c & d \end{vmatrix} = (ad - cb) \ .$$

THM. Equivalency Theorem: For an $n \times n$ matrix A, the following properties are equivalent:

- The inverse of A exists.
- The determinate of A is nonzero.
- The rows of A form a basis for $\mathbf{R}^n$.
- The columns of A form a basis for $\mathbf{R}^n$.

- As a map from $\mathbf{R}^n$ to $\mathbf{R}^n$, A is injective (one-to-one).
- As a map from $\mathbf{R}^n$ to $\mathbf{R}^n$, A is surjective (onto).
- The equation $Ax = 0$ implies $x = 0$.
- For each $b \in \mathbf{R}^n$, there is exactly one $x \in \mathbf{R}^n$ such that $Ax = b$.
- A is a product of elementary matrices.

Representing a Linear System Using Matrix Notation

Given an $n \times n$ system of equations, the system can be reexpressed in terms of a $n \times n$ matrix (A) of coefficients, a $n \times 1$ matrix (b) of constants on the right and a $n \times 1$ matrix of unknowns (x), i.e.,

$$Ax = b \ .$$

Consider the following systems of equations,

$$\begin{aligned} ax + by &= f_1 \\ cx + dy &= f_2 \,. \end{aligned}$$

In matrix notation, the above system is given by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$$

If the inverse to A exists, then a unique solution exists and the solutions is given by $x = A^{-1}b$. It can be shown that for a 2×2 matrix A, if it exists the inverse is given by

$$A^{-1} = \frac{1}{ad - cb} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix}$$

so that the solution for our example is

$$\begin{bmatrix} x \\ y \end{bmatrix} = \frac{1}{ab - cd} \begin{bmatrix} d & -b \\ -c & a \end{bmatrix} \begin{bmatrix} f_1 \\ f_2 \end{bmatrix} = \frac{1}{ab - cd} \begin{bmatrix} d \cdot f_1 - b \cdot f_2 \\ -c \cdot f_1 + a \cdot f_2 \end{bmatrix} \ .$$

A Fundamental Theorem About Linear System

THM. Equivalence of Linear Systems: If one system of equations is obtained from another by a finite sequence of operations (adding two equations, multiplying an equation by a constant, adding a multiple of one equation to another), then the two systems are equivalent (THIS IS VERY IMPORTANT TO EVERYTHING WE DO FROM HERE ON IN). Example, the following systems are equivalent,

$$\begin{aligned} ax \,+\, by &= f_1 \\ cx \,+\, dy &= f_2\,. \end{aligned}$$

and

$$\begin{aligned} ax \,+\, by &= f_1 \\ (c-\lambda a)x \,+\, (d-\lambda b)y &= (f_2-\lambda f_1) \end{aligned}$$

where system two was obtained by adding $-\lambda$ times equation one to equation two.

What the Above Fundamental Theorem Implies about Matrix Formulation

In terms of matrices, the above two systems are given by

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$$

and

$$\begin{bmatrix} a & b \\ (c-\lambda a) & (d-\lambda b) \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f_1 \\ (f_2-\lambda f_1) \end{bmatrix}\,.$$

We see that applying elementary operations to a linear system of equations is equivalent to applying these operations to the appropriate rows of the matrix A and b. So we can work with just the coefficients of the matrices and neglect the vector of unknowns.

Def. Augmented Matrix: Given the system of equations $Ax = b$, we define the augmented matrix to be $[A|b]$. For our example

$$\begin{bmatrix} a & b \\ c & d \end{bmatrix}\begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} f_1 \\ f_2 \end{bmatrix}$$

the augmented matrix for the system is

$$\left[\begin{array}{cc|c} a & b & f_1 \\ c & d & f_2 \end{array}\right]\,.$$

Labelling Elementary Operations for an Augmented System

1. Interchanging rows s and t of matrix A

$$A : A_s \leftrightarrow A_t$$

2. Multiplying row s of matrix a by a non-zero constant λ

$$A : A_s \rightarrow \lambda A_s$$

3. Replacing row s by adding a multiple of row t to row s

$$A : A_s \rightarrow A_s + \lambda A_t$$

Example: consider the system of equations

$$\begin{aligned} 1x &+ 3y &= 7 \\ 2x &+ 5y &= 2\,. \end{aligned}$$

Interchange rows one and two

$$\left[\begin{array}{cc|c} 1 & 3 & 7 \\ 2 & 5 & 2 \end{array}\right] \overset{A : A_1 \leftrightarrow A_2}{\Rightarrow} \left[\begin{array}{cc|c} 2 & 5 & 2 \\ 1 & 3 & 7 \end{array}\right].$$

Adding $\frac{-1}{2}$ row one to row two

$$\left[\begin{array}{cc|c} 2 & 5 & 2 \\ 1 & 3 & 7 \end{array}\right] \overset{A : A_2 \rightarrow A_2 + \frac{-1}{2} A_1}{\Rightarrow} \left[\begin{array}{cc|c} 2 & 5 & 2 \\ 1-1 & 3+\frac{-5}{2} & 7-1 \end{array}\right] = \left[\begin{array}{cc|c} 2 & 5 & 2 \\ 0 & \frac{1}{2} & 6 \end{array}\right].$$

Multiplying row two by 2.

$$\left[\begin{array}{cc|c} 2 & 5 & 2 \\ 0 & \frac{1}{2} & 6 \end{array}\right] \overset{A : A_2 \rightarrow 2A_2}{\Rightarrow} \left[\begin{array}{cc|c} 2 & 5 & 2 \\ 0 & 1 & 12 \end{array}\right].$$

Adding -5 row two to row one

$$\left[\begin{array}{cc|c} 2 & 5 & 2 \\ 0 & 1 & 12 \end{array}\right] \overset{A : A_1 \rightarrow A_1 + (-5) A_2}{\Rightarrow} \left[\begin{array}{cc|c} 2 & 0 & -58 \\ 0 & 1 & 12 \end{array}\right].$$

Multiplying row one by $\frac{1}{2}$

$$\left[\begin{array}{cc|c} 2 & 0 & -58 \\ 0 & 1 & 12 \end{array}\right] \overset{A: A_1 \to \frac{1}{2}A_1}{\Rightarrow} \left[\begin{array}{cc|c} 1 & 0 & -29 \\ 0 & 1 & 12 \end{array}\right].$$