IB Paper 7: Linear Algebra   Handout 2

2.6 The Column Picture for Matrix Multiplication, revisited

We saw in Handout 1 that when we multiply *a matrix and a vector*, we get

$$\mathbf{A}\,\mathbf{x} \;=\; \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix} \;=\; \begin{bmatrix} \uparrow & \uparrow & \dots & \uparrow \\ \underline{a}_1 & \underline{a}_2 & & \underline{a}_n \\ \downarrow & \downarrow & \dots & \downarrow \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \dots \\ x_n \end{bmatrix}$$

$$= \; x_1\underline{a}_1 + x_2\underline{a}_2 + x_3\underline{a}_3 + \dots + x_n\underline{a}_n$$

In this section we will show that something similar happens when we multiply two *matrices* together.

Consider the simplest case of the product of two simple matrices

$$\mathbf{A} \;=\; \mathbf{BC} \;=\; \begin{bmatrix} a & b \\ c & d \end{bmatrix} \begin{bmatrix} e & f \\ g & h \end{bmatrix} \;=\; \begin{bmatrix} ae+bg & af+bh \\ ce+dg & cf+gh \end{bmatrix} \;=\; \begin{bmatrix} ae & af \\ ce & cf \end{bmatrix} + \begin{bmatrix} bg & bh \\ dg & dh \end{bmatrix}$$

This pattern generalises

$$\mathbf{A} \;=\; \mathbf{BC} \;=\; \begin{bmatrix} b_{11} & b_{12} & b_{13} \\ b_{21} & b_{22} & b_{23} \\ b_{31} & b_{32} & b_{33} \end{bmatrix} \begin{bmatrix} c_{11} & c_{12} & c_{13} \\ c_{21} & c_{22} & c_{23} \\ c_{31} & c_{32} & c_{33} \end{bmatrix}$$

$$= \begin{bmatrix} b_{11}c_{11}+b_{12}c_{21}+b_{13}c_{31} & b_{11}c_{12}+b_{12}c_{22}+b_{13}c_{32} & b_{11}c_{13}+b_{12}c_{23}+b_{13}c_{33} \\ b_{21}c_{11}+b_{22}c_{21}+b_{23}c_{31} & b_{21}c_{12}+b_{22}c_{22}+b_{23}c_{32} & b_{21}c_{13}+b_{22}c_{23}+b_{23}c_{33} \\ b_{31}c_{11}+b_{32}c_{21}+b_{33}c_{31} & b_{31}c_{12}+b_{32}c_{22}+b_{33}c_{32} & b_{31}c_{13}+b_{32}c_{23}+b_{33}c_{33} \end{bmatrix}$$

Splitting into 3 matrices

$$= \begin{bmatrix} b_{11}c_{11} & b_{11}c_{12} & b_{11}c_{13} \\ b_{21}c_{11} & b_{21}c_{12} & b_{21}c_{13} \\ b_{31}c_{11} & b_{31}c_{12} & b_{31}c_{13} \end{bmatrix} + \begin{bmatrix} b_{12}c_{21} & b_{12}c_{22} & b_{12}c_{23} \\ b_{22}c_{21} & b_{22}c_{22} & b_{22}c_{33} \\ b_{32}c_{21} & b_{32}c_{22} & b_{32}c_{23} \end{bmatrix} + \begin{bmatrix} b_{13}c_{31} & b_{13}c_{32} & b_{13}c_{33} \\ b_{23}c_{31} & b_{23}c_{32} & b_{23}c_{33} \\ b_{33}c_{31} & b_{33}c_{32} & b_{33}c_{33} \end{bmatrix}$$

Now, these matrices can be recognised as

If we denote the *columns* of **B** as $\underline{b}_1$, $\underline{b}_2$, etc. and the *rows* of **C** as $\tilde{\underline{c}}_1$, $\tilde{\underline{c}}_2$, etc. then this can be written more succinctly as

A product of vectors $\underline{x}\,\underline{y}^{\mathrm{T}}$ is referred to as an *outer product* of $\underline{x}$ and $\underline{y}$. (the dot product $\underline{x}^{\mathrm{T}}\,\underline{y} = \underline{x}\,\boldsymbol{.}\,\underline{y}$ is also called an *inner product*). So that

> **A** = **BC** = sum of the outer products of each of the columns of **B** with the corresponding row of **C**.

This is not the only consequence of viewing the product like this. If we denote the columns of **A** as $\underline{a}_1$, $\underline{a}_2$, etc.

$$\begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \underline{a}_1 & \underline{a}_2 & \underline{a}_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow \\ \underline{b}_1 \\ \downarrow \end{bmatrix}\begin{bmatrix} c_{11} & c_{12} & c_{13} \end{bmatrix} + \begin{bmatrix} \uparrow \\ \underline{b}_2 \\ \downarrow \end{bmatrix}\begin{bmatrix} c_{21} & c_{22} & c_{23} \end{bmatrix} + \begin{bmatrix} \uparrow \\ \underline{b}_3 \\ \downarrow \end{bmatrix}\begin{bmatrix} c_{31} & c_{32} & c_{33} \end{bmatrix}$$

is equivalent to

$$\begin{bmatrix} \uparrow \\ \underline{a}_1 \\ \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow \\ \underline{b}_1 \\ \downarrow \end{bmatrix}\begin{bmatrix} c_{11} \end{bmatrix} + \begin{bmatrix} \uparrow \\ \underline{b}_2 \\ \downarrow \end{bmatrix}\begin{bmatrix} c_{21} \end{bmatrix} + \begin{bmatrix} \uparrow \\ \underline{b}_3 \\ \downarrow \end{bmatrix}\begin{bmatrix} c_{31} \end{bmatrix} \qquad \begin{bmatrix} \uparrow \\ \underline{a}_2 \\ \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow \\ \underline{b}_1 \\ \downarrow \end{bmatrix}\begin{bmatrix} c_{12} \end{bmatrix} + \begin{bmatrix} \uparrow \\ \underline{b}_2 \\ \downarrow \end{bmatrix}\begin{bmatrix} c_{22} \end{bmatrix} + \begin{bmatrix} \uparrow \\ \underline{b}_3 \\ \downarrow \end{bmatrix}\begin{bmatrix} c_{32} \end{bmatrix}$$

and

$$\begin{bmatrix} \uparrow \\ \underline{a}_3 \\ \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow \\ \underline{b}_1 \\ \downarrow \end{bmatrix}\begin{bmatrix} c_{13} \end{bmatrix} + \begin{bmatrix} \uparrow \\ \underline{b}_2 \\ \downarrow \end{bmatrix}\begin{bmatrix} c_{23} \end{bmatrix} + \begin{bmatrix} \uparrow \\ \underline{b}_3 \\ \downarrow \end{bmatrix}\begin{bmatrix} c_{33} \end{bmatrix}$$

Now, there is nothing special about 3×3 matrices. For the general case of

$$\mathbf{A} = \mathbf{B}\,\mathbf{C}$$

where **A** is an $m \times n$ matrix, **B** is an $m \times k$ matrix, **C** is an $k \times n$ matrix,

$$\begin{bmatrix} \uparrow & \uparrow & \cdots & \uparrow \\ \underline{a}_1 & \underline{a}_2 & \cdots & \underline{a}_n \\ \downarrow & \downarrow & \cdots & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow \\ \underline{b}_1 \\ \downarrow \end{bmatrix}\begin{bmatrix} c_{11} & c_{12} & \cdots & c_{1n} \end{bmatrix} + \begin{bmatrix} \uparrow \\ \underline{b}_2 \\ \downarrow \end{bmatrix}\begin{bmatrix} c_{21} & c_{22} & \cdots & c_{2n} \end{bmatrix} + \cdots + \begin{bmatrix} \uparrow \\ \underline{b}_k \\ \downarrow \end{bmatrix}\begin{bmatrix} c_{k1} & c_{k2} & \cdots & c_{kn} \end{bmatrix} \qquad (2.1)$$

or more succinctly

This is a really powerful relationship. It enables us to move back and forth between matrix multiplication and relationships between sets of vectors. For the j'th column of $\mathbf{A}$

$$\begin{bmatrix} \uparrow \\ \underline{a}_j \\ \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow \\ \underline{b}_1 \\ \downarrow \end{bmatrix} \begin{bmatrix} c_{1j} \end{bmatrix} + \begin{bmatrix} \uparrow \\ \underline{b}_2 \\ \downarrow \end{bmatrix} \begin{bmatrix} c_{2j} \end{bmatrix} + \dots + \begin{bmatrix} \uparrow \\ \underline{b}_k \\ \downarrow \end{bmatrix} \begin{bmatrix} c_{kj} \end{bmatrix}$$

So, for example,

$$\begin{bmatrix} 3 & 7 & 0 \\ 0 & 5 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \\ 1 & 4 & -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \begin{bmatrix} 1 & -1 & 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix} \begin{bmatrix} 1 & 4 & -1 \end{bmatrix}$$

is equivalent to

$$\begin{bmatrix} 3 \\ 0 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} 1 + \begin{bmatrix} 2 \\ 1 \end{bmatrix} 1 \qquad \begin{bmatrix} 7 \\ 5 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad + \begin{bmatrix} 2 \\ 1 \end{bmatrix} \qquad \begin{bmatrix} 0 \\ -3 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \qquad + \begin{bmatrix} 2 \\ 1 \end{bmatrix}$$

Note that we have started to write the scalars multiplying the vectors *after* them. In the current framework, we are thinking of scalars as 1×1 matrices and they have to go behind since, when $n > 1$, you can multiply a $(n \times 1)$ matrix times a (1×1) but not a (1×1) times a $(n \times 1)$.

Examples

We can also use this to find matrices that manipulate column vectors.

(a) Find the matrix which swaps the 2nd and 3rd columns of a 3×3 matrix.

If we write this as $\mathbf{A} = \mathbf{B}\,\mathbf{C}$, then, in terms of the columns of $\mathbf{A}$ and $\mathbf{B}$, this is equivalent to

$$\underline{a}_1 = \underline{b}_1 \qquad \underline{a}_2 = \underline{b}_3 \qquad \underline{a}_3 = \underline{b}_2$$

i.e.

$$\begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \underline{a}_1 & \underline{a}_2 & \underline{a}_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow \\ \underline{b}_1 \\ \downarrow \end{bmatrix} \qquad + \begin{bmatrix} \uparrow \\ \underline{b}_2 \\ \downarrow \end{bmatrix} \qquad + \begin{bmatrix} \uparrow \\ \underline{b}_3 \\ \downarrow \end{bmatrix}$$

$$= \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \underline{b}_1 & \underline{b}_2 & \underline{b}_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix} \begin{bmatrix} \quad & \quad & \quad \\ \quad & \quad & \quad \\ \quad & \quad & \quad \end{bmatrix}$$

C is a *permutation matrix*, that shuffles the rows. Note that its columns are mutually orthogonal unit vectors, i.e. it is an orthogonal matrix meaning that its inverse is its transpose.

(b) Find the matrix which subtracts 1 × first column and 2 × second column from the 3rd, while leaving the first and second columns alone.

In terms of column vectors

$$\begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \underline{a}_1 & \underline{a}_2 & \underline{a}_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow \\ \underline{b}_1 \\ \downarrow \end{bmatrix} \qquad + \begin{bmatrix} \uparrow \\ \underline{b}_2 \\ \downarrow \end{bmatrix} \qquad + \begin{bmatrix} \uparrow \\ \underline{b}_3 \\ \downarrow \end{bmatrix}$$

$$= \begin{bmatrix} \uparrow & \uparrow & \uparrow \\ \underline{b}_1 & \underline{b}_2 & \underline{b}_3 \\ \downarrow & \downarrow & \downarrow \end{bmatrix} \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}$$

2.7 The Row Picture for Matrix Multiplication

It is usually true for matrices that, if there is a property that applies to columns, then there is a corresponding one that applies to rows. Returning to (2.1) & (2.2) for the general case

$$\mathbf{A} = \mathbf{B}\,\mathbf{C}$$

where **A** is an $m \times n$ matrix, **B** is an $m \times k$ matrix, **C** is an $k \times n$ matrix, we see that

$$\mathbf{A} = \underline{b}_1 \underline{\tilde{c}}_1^{\mathrm{T}} + \underline{b}_2 \underline{\tilde{c}}_2^{\mathrm{T}} + \underline{b}_3 \underline{\tilde{c}}_3^{\mathrm{T}} + \dots + \underline{b}_k \underline{\tilde{c}}_k^{\mathrm{T}}$$

can also be interpreted as

$$\begin{bmatrix} \leftarrow & \underline{\tilde{a}}_1 & \rightarrow \\ \leftarrow & \underline{\tilde{a}}_2 & \rightarrow \\ \dots & \dots & \dots \\ \leftarrow & \underline{\tilde{a}}_m & \rightarrow \end{bmatrix} = \begin{bmatrix} b_{11} \\ b_{21} \\ \dots \\ b_{m1} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_1 & \rightarrow \end{bmatrix} + \begin{bmatrix} b_{12} \\ b_{22} \\ \dots \\ b_{m2} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_2 & \rightarrow \end{bmatrix} + \dots + \begin{bmatrix} b_{1k} \\ b_{2k} \\ \dots \\ b_{mk} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_k & \rightarrow \end{bmatrix}$$

i.e.

$$\begin{bmatrix} \leftarrow & \underline{\tilde{a}}_1 & \rightarrow \end{bmatrix} = \begin{bmatrix} b_{11} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_1 & \rightarrow \end{bmatrix} + \begin{bmatrix} b_{12} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_2 & \rightarrow \end{bmatrix} + \dots + \begin{bmatrix} b_{1k} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_k & \rightarrow \end{bmatrix}$$

$$\begin{bmatrix} \leftarrow & \underline{\tilde{a}}_2 & \rightarrow \end{bmatrix} = \begin{bmatrix} b_{21} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_1 & \rightarrow \end{bmatrix} + \begin{bmatrix} b_{22} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_2 & \rightarrow \end{bmatrix} + \dots + \begin{bmatrix} b_{2k} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_k & \rightarrow \end{bmatrix}$$

$$\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots\dots$$

$$\begin{bmatrix} \leftarrow & \underline{\tilde{a}}_m & \rightarrow \end{bmatrix} = \begin{bmatrix} b_{m1} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_1 & \rightarrow \end{bmatrix} + \begin{bmatrix} b_{m2} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_2 & \rightarrow \end{bmatrix} + \dots + \begin{bmatrix} b_{mk} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_k & \rightarrow \end{bmatrix}$$

where the notation $\underline{\tilde{c}}_1$ has been used to indicate that we are talking about vectors which represent the *rows*, rather than the columns. The general form for the i'th row is

$$\begin{bmatrix} \leftarrow & \underline{\tilde{a}}_i & \rightarrow \end{bmatrix} = \begin{bmatrix} b_{i1} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_1 & \rightarrow \end{bmatrix} + \begin{bmatrix} b_{i2} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_2 & \rightarrow \end{bmatrix} + \dots + \begin{bmatrix} b_{ik} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_k & \rightarrow \end{bmatrix}$$

So, for example,

$$\begin{bmatrix} 3 & 7 & 0 \\ 0 & 5 & -3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ -1 & 1 \end{bmatrix}\begin{bmatrix} 1 & -1 & 2 \\ 1 & 4 & -1 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix}\begin{bmatrix} 1 & -1 & 2 \end{bmatrix} + \begin{bmatrix} 2 \\ 1 \end{bmatrix}\begin{bmatrix} 1 & 4 & -1 \end{bmatrix}$$

is equivalent to

$$\begin{bmatrix} 3 & 7 & 0 \end{bmatrix} =$$

and

$$\begin{bmatrix} 0 & 5 & -3 \end{bmatrix} =$$

Examples

(a) Find the matrix which subtracts 1 × first row and 2 × second row from the 3rd, while leaving the first and second rows alone.

In terms of row vectors

$$\underset{\sim}{a}_1 = \underset{\sim}{c}_1 \qquad \underset{\sim}{a}_2 = \underset{\sim}{c}_2 \qquad \underset{\sim}{a}_3 = \underset{\sim}{c}_3 - \underset{\sim}{c}_1 - 2\underset{\sim}{c}_2$$

$$\Rightarrow \qquad \begin{bmatrix} \leftarrow & \underset{\sim}{a}_1 & \rightarrow \\ \leftarrow & \underset{\sim}{a}_2 & \rightarrow \\ \leftarrow & \underset{\sim}{a}_3 & \rightarrow \end{bmatrix} = \begin{bmatrix} \qquad & \qquad & \qquad \\ & & \\ & & \end{bmatrix} \begin{bmatrix} \leftarrow & \underset{\sim}{c}_1 & \rightarrow \\ \leftarrow & \underset{\sim}{c}_2 & \rightarrow \\ \leftarrow & \underset{\sim}{c}_3 & \rightarrow \end{bmatrix}$$

(b) Find the matrix which cyclically permutes the rows of a 4 × 4 matrix (i.e. 1 → 2, 2 → 3, 3 → 4, 4 → 1

$$\Rightarrow \qquad \begin{bmatrix} \leftarrow & \underset{\sim}{a}_1 & \rightarrow \\ \leftarrow & \underset{\sim}{a}_2 & \rightarrow \\ \leftarrow & \underset{\sim}{a}_3 & \rightarrow \\ \leftarrow & \underset{\sim}{a}_4 & \rightarrow \end{bmatrix} = \begin{bmatrix} \qquad & \qquad & \qquad & \qquad \\ & & & \\ & & & \\ & & & \end{bmatrix} \begin{bmatrix} \leftarrow & \underset{\sim}{c}_1 & \rightarrow \\ \leftarrow & \underset{\sim}{c}_2 & \rightarrow \\ \leftarrow & \underset{\sim}{c}_3 & \rightarrow \\ \leftarrow & \underset{\sim}{c}_4 & \rightarrow \end{bmatrix}$$

But hang on a minute

Because addition of matrices can be done in any order, the relationship

$$\mathbf{A} = \underline{b}_1 \underset{\sim}{c}_1^{\mathrm{T}} + \underline{b}_2 \underset{\sim}{c}_2^{\mathrm{T}} + \underline{b}_3 \underset{\sim}{c}_3^{\mathrm{T}} = \underline{b}_3 \underset{\sim}{c}_3^{\mathrm{T}} + \underline{b}_1 \underset{\sim}{c}_1^{\mathrm{T}} + \underline{b}_2 \underset{\sim}{c}_2^{\mathrm{T}}$$

appears to suggest that we can re-arrange the columns of **B** provided we apply the same re-arrangement to the rows of **C**. This is in fact true. If the permutation matrix to rearrange the columns of **B** is **P** (as a *post*-multiplier), then the permutation matrix for the rows of **C** is $\mathbf{P}^{\mathrm{T}}$ (as a *pre*-multiplier).

3. LU Factorisation

This technique is inspired by traditional (= Gaussian elimination) methods for solving simultaneous equations. The idea is that we factorise an $m \times n$ matrix $\mathbf{A}$ into the form

$$\mathbf{A} = \mathbf{LU}$$

where $\mathbf{L}$ is lower-triangular $m \times m$ matrix with 1's down the leading diagonal and $\mathbf{U}$ is an upper-echelon matrix which is the same shape as $\mathbf{A}$.

Lower triangular means that $\mathbf{L}$ has non-zero terms only on and below the leading diagonal (e.g 4×4)

$$\mathbf{L} = \begin{bmatrix} 1 & 0 & 0 & 0 \\ * & 1 & 0 & 0 \\ * & * & 1 & 0 \\ * & * & * & 1 \end{bmatrix} \quad \text{where * represents a possibly non-zero value.}$$

Upper echelon means that all non-zero elements are on or above the leading diagonal. e.g. (3×4)

$$\mathbf{U} = \begin{bmatrix} * & * & * & * \\ 0 & * & * & * \\ 0 & 0 & * & * \end{bmatrix} \quad \text{where * represents a possibly non-zero value.}$$

If $\mathbf{A}$ is square, then so is $\mathbf{U}$ and then $\mathbf{U}$ is said to be upper triangular.

The technique is best demonstrated by examples. We are seeking to write

$$\mathbf{A} = \mathbf{LU} = \underline{l}_1 \underline{\tilde{u}}_1^{\mathrm{T}} + \underline{l}_2 \underline{\tilde{u}}_2^{\mathrm{T}} + \underline{l}_3 \underline{\tilde{u}}_3^{\mathrm{T}} + \dots + \underline{l}_m \underline{\tilde{u}}_m^{\mathrm{T}}$$

and $\underline{l}_1, \underline{l}_2, \dots$ are the columns of $\mathbf{L}$ and $\underline{\tilde{u}}_1, \underline{\tilde{u}}_2, \dots$ are the rows of $\mathbf{U}$. These rows and columns can be generated in turn because of the particular patterns of 1's and 0's.

$$\underset{\mathbf{A}}{\begin{bmatrix} 2 & 1 & -2 \\ 6 & 4 & -3 \\ 4 & 3 & 0 \end{bmatrix}} = \underset{\underline{l}_1 \underline{\tilde{u}}_1^{\mathrm{T}}}{\begin{bmatrix} \quad & \quad & \quad \\ \quad & \quad & \quad \\ \quad & \quad & \quad \end{bmatrix}} \underset{+}{+} \underset{\text{rest}}{\begin{bmatrix} \quad & \quad & \quad \\ \quad & \quad & \quad \\ \quad & \quad & \quad \end{bmatrix}}$$

We have chosen $\underline{\tilde{u}}_1$ to be the first row of $\mathbf{A}$ and $\underline{l}_1$ is determined by 1 as the first component and then whatever is necessary to make the first column of $\underline{l}_1 \underline{\tilde{u}}_1^{\mathrm{T}}$ equal to the first column of $\mathbf{A}$. These choices give us a zero top row and a zero first column in the matrix representing "the rest". i.e.

$$\mathbf{Rest} = \mathbf{A} - \underline{l}_1 \underline{\tilde{u}}_1^{\mathrm{T}}$$

We now repeat the exercise with the “rest” to choose $\underline{l}_2$ and $\underline{\tilde{u}}_2$

$$\begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} = \begin{matrix}0\\1\\1\end{matrix} \begin{array}{c} \begin{matrix}0 & 1 & 3\end{matrix} \\ \begin{bmatrix} & & \\ & & \\ & & \end{bmatrix} \end{array} + \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Rest $\qquad\qquad \underline{l}_2\ \tilde{\underline{u}}_2^{\mathrm{T}} \quad + \quad$ rest = $\underline{l}_3\ \tilde{\underline{u}}_3^{\mathrm{T}}$

and, in general, repeat. In this case we can, however, spot $\underline{l}_3$ and $\tilde{\underline{u}}_3$ immediately.

Note that, after each stage of the process, "rest" has one more zero row and one more zero column.

Gathering results

$$\mathbf{A} = \begin{bmatrix} \\ \\ \end{bmatrix}\begin{bmatrix} & & \end{bmatrix} + \begin{bmatrix} \\ \\ \end{bmatrix}\begin{bmatrix} & & \end{bmatrix} + \begin{bmatrix} \\ \\ \end{bmatrix}\begin{bmatrix} & & \end{bmatrix}$$

$$= \begin{bmatrix} 1 & 0 & 0 \\ & 1 & 0 \\ & & 1 \end{bmatrix} \begin{bmatrix} & & \\ 0 & & \\ 0 & 0 & 1 \end{bmatrix} = \mathbf{L}\,\mathbf{U}$$

3.2 A Second Example

This time I have already inserted the numbers which are determined by the chosen 1's and 0's in **L** and **U**.

$$\underset{\mathbf{A}}{\begin{bmatrix} 1 & 2 & 0 \\ -1 & -1 & 3 \\ 3 & 7 & 5 \end{bmatrix}} = \begin{matrix}1\\ \\ \end{matrix} \underset{\underline{l}_1\ \tilde{\underline{u}}_1^{\mathrm{T}}}{\begin{bmatrix} & & \\ & & \\ & & \end{bmatrix}} + \underset{\text{rest}}{\begin{bmatrix} 0 & 0 & 0 \\ 0 & & \\ 0 & & \end{bmatrix}}$$

$$\begin{matrix}0\\1\\ \end{matrix} \begin{array}{c} \begin{matrix}0 & & \end{matrix} \\ \begin{bmatrix} 0 & 0 & 0 \\ 0 & & \\ 0 & & \end{bmatrix} \end{array} + \begin{matrix}0\\0\\1\end{matrix} \begin{array}{c} \begin{matrix}0 & 0 & \end{matrix} \\ \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & \end{bmatrix} \end{array}$$

$$\mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ & 1 & 0 \\ & & 1 \end{bmatrix} \begin{bmatrix} & & \\ 0 & & \\ 0 & 0 & \end{bmatrix}$$

3.3 Non-square matrices

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 6 & 3 & 12 \\ -1 & 2 & 2 & 8 \end{bmatrix}$$

No Problem.

$$\begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 6 & 3 & 12 \\ -1 & 2 & 2 & 8 \end{bmatrix} = 1 \begin{bmatrix} & & & \\ & & & \\ & & & \end{bmatrix} + \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & & \end{bmatrix}$$

$$0 \qquad\qquad\qquad\qquad 0 \quad 0$$

$$\begin{matrix} 0 \\ 1 \\ \end{matrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & & & \\ 0 & & & \end{bmatrix} + \begin{matrix} 0 \\ 0 \\ 1 \end{matrix} \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & & \end{bmatrix}$$

$$\Rightarrow \qquad \mathbf{A} = \begin{bmatrix} 1 & 0 & 0 \\ & 1 & 0 \\ & & 1 \end{bmatrix} \begin{bmatrix} & & & \\ 0 & & & \\ 0 & 0 & & \end{bmatrix}$$

3.4 A first look at the solution of A$\underline{x}$ = $\underline{b}$

Consider the equations

$$\begin{aligned} 2x + y - 2z &= 1 \\ 6x + 4y - 3z &= 5 \\ 4x + 3y &= 5 \end{aligned} \qquad \Rightarrow \mathbf{A} = \begin{bmatrix} 2 & 1 & -2 \\ 6 & 4 & -3 \\ 4 & 3 & 0 \end{bmatrix} \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix}$$

We have already performed LU decomposition on **A** and we shall see how the simple forms for **L** and **U** pay off when we try to solve these equations.

$$\mathbf{A} = \mathbf{LU} \quad \text{where} \quad \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix} \quad \mathbf{U} = \begin{bmatrix} 2 & 1 & -2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix}$$

The solution method then is simply

$$\mathbf{A}\underline{x} = (\mathbf{LU})\underline{x} = \mathbf{L}(\mathbf{U}\underline{x}) = \underline{b}$$

We solve this in two steps,

(i) find $\underline{c}$ from $\mathbf{L}\underline{c} = \underline{b}$

and then (ii)find $\underline{x}$ from $\mathbf{U}\underline{x} = \underline{c}$.

$$\text{i.e. } \mathbf{L}\underline{c} = \underline{b} \quad \Rightarrow \quad \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix} \begin{bmatrix} c_1 \\ c_2 \\ c_3 \end{bmatrix} = \begin{bmatrix} 1 \\ 5 \\ 5 \end{bmatrix}$$

Solve by *forward* substitution (i.e. solve the first equation, substitute into the second, etc.)

Then

$$\mathbf{U}\underline{x} = \underline{c} \quad \Rightarrow \quad \begin{bmatrix} 2 & 1 & -2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 1 \end{bmatrix}$$

Solve by *backward* substitution (i.e. solve the last equation, substitute into the second last, etc.)

A bit of jargon: The first non-zero elements of **U** in each row are called *pivots*. The pivot for x is 2, the pivot for y is 1, that for z is 1. At any stage in the back substitution: *pivot* × *unknown* = *known*.

The terms below the diagonal of **L** are *multipliers*. 3 and 2 are the multipliers for x. 1 is the multiplier for y.

$$\mathbf{U} = \begin{bmatrix} 2 & 1 & -2 \\ 0 & 1 & 3 \\ 0 & 0 & 1 \end{bmatrix} \qquad \mathbf{L} = \begin{bmatrix} 1 & 0 & 0 \\ 3 & 1 & 0 \\ 2 & 1 & 1 \end{bmatrix}$$

We shall discuss in a later section that large pivots and the small multipliers keeps the problem well-conditioned. We will extend the LU decomposition to ensure that this happens.

You can now do Examples Paper 1 Questions 4 and 5

Key Points from Lecture

- $\mathbf{A} = \mathbf{B}\,\mathbf{C}$ $\quad$ ($\mathbf{A}\ m \times n$ $\ \mathbf{B}\ m \times k$ $\ \mathbf{C}\ k \times n$)

is equivalent to

$$\mathbf{A} = \underline{b}_1 \underline{\tilde{c}}_1^{\mathrm{T}} + \underline{b}_2 \underline{\tilde{c}}_2^{\mathrm{T}} + \underline{b}_3 \underline{\tilde{c}}_3^{\mathrm{T}} + \dots + \underline{b}_k \underline{\tilde{c}}_k^{\mathrm{T}}$$

which is a sum of outer products.

- This is a relationship between the columns of **A** and the columns of **B** with the elements of **C** as the coefficients.

$$\begin{bmatrix} \uparrow \\ \underline{a}_j \\ \downarrow \end{bmatrix} = \begin{bmatrix} \uparrow \\ \underline{b}_1 \\ \downarrow \end{bmatrix} \begin{bmatrix} c_{1j} \end{bmatrix} + \begin{bmatrix} \uparrow \\ \underline{b}_2 \\ \downarrow \end{bmatrix} \begin{bmatrix} c_{2j} \end{bmatrix} + \dots + \begin{bmatrix} \uparrow \\ \underline{b}_k \\ \downarrow \end{bmatrix} \begin{bmatrix} c_{kj} \end{bmatrix}$$

- In addition, it also represents a relationship between the rows of **A** and the rows of **C** with the elements of **B** representing the coefficients.

$$\begin{bmatrix} \leftarrow & \underline{\tilde{a}}_i & \rightarrow \end{bmatrix} = \begin{bmatrix} b_{i1} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_1 & \rightarrow \end{bmatrix} + \begin{bmatrix} b_{i2} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_2 & \rightarrow \end{bmatrix} + \dots + \begin{bmatrix} b_{ik} \end{bmatrix} \begin{bmatrix} \leftarrow & \underline{\tilde{c}}_k & \rightarrow \end{bmatrix}$$

LU Factorisation

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 1 & 3 \\ 2 & 6 & 3 & 12 \\ -1 & 2 & 2 & 8 \end{bmatrix} = \underset{\mathbf{L}}{\begin{bmatrix} 1 & 0 & 0 \\ 2 & 1 & 0 \\ -1 & 2 & 1 \end{bmatrix}} \underset{\mathbf{U}}{\begin{bmatrix} 1 & 2 & 1 & 3 \\ 0 & 2 & 1 & 6 \\ 0 & 0 & 1 & -1 \end{bmatrix}}$$

L is always a square matrix, while **U** always has the same shape as **A.**

Solution of A$\underline{x}$ = $\underline{b}$

Recast this as (i) find $\underline{c}$ from $\mathbf{L}\underline{c} = \underline{b}$

and then (ii)find $\underline{x}$ from $\mathbf{U}\underline{x} = \underline{c}$.