CSci 5304, F'24 Solution keys to some exercises from: Set 1

✍2 $(A^T)^T = ??$ **Solution:** $(A^T)^T = A$

✍3 $(AB)^T = ??$ **Solution:** $(AB)^T = B^T A^T$

✍4 $(A^H)^H = ??$ **Solution:** $(A^H)^H = A$

✍5 $(A^H)^T = ??$ **Solution:** $(A^H)^T = \bar{A}$

✍6 $(ABC)^T = ??$ **Solution:** $(ABC)^T = C^T B^T A^T$

✍7 True/False: $(AB)C = A(BC)$ **Solution:** $\rightarrow$ True

✍8 True/False: $AB = BA$ **Solution:** $\rightarrow$ false in general

✍9 True/False: $AA^T = A^T A$ **Solution:** $\rightarrow$ false in general

✍12 Complexity? [number of multiplications and additions for matrix multiply]

Solution: Let $A \in \mathbb{R}^{m \times n}$ and $B \in \mathbb{R}^{n \times p}$. Then the product AB requires $2mnp$ operations (there are mp entries in all and each of them requires $2n$ operations). □

✍13 What happens to these 3 different approches to matrix-matrix multiplication when B has one column ($p = 1$)?

Solution: In the first: $C_{:,j}$ the j=th column of C is a linear combination of the columns of A. This is the usual matrix-vector product.

In the second: $C_{i,:}$ is just a number which is the inner product of the ith row of A with the column B.

The 3rd formula will give the exact same expression as the first. $\square$

✍14 Characterize the matrices AA^T and A^TA when A is of dimension $n \times 1$.

Solution: When $A \in \mathbb{R}^{n\times 1}$ then AA^T is a rank-one $n \times n$ matrix and A^TA is a scalar: the inner product of the column A with itself. $\square$

✍15 Show that for 2 vectors u, v we have $v^T \otimes u = uv^T$

Solution: The $j-$th column of $v^T \otimes u$ is just $v_j.u$ This is also the jth column of uv^T. $\square$

✍16 Show that $A \in \mathbb{R}^{m\times n}$ is of rank one iff [if and only if] there exist two nonzero vectors $\boldsymbol{u} \in \mathbb{R}^m$ and $\boldsymbol{v} \in \mathbb{R}^n$ such that

$$\boldsymbol{A} = \boldsymbol{u}\boldsymbol{v}^T.$$

What are the eigenvalues and eigenvectors of $\boldsymbol{A}$?

Solution: (a: First part)

$\leftarrow$ First we show that: When both $\boldsymbol{u}$ and $\boldsymbol{v}$ are nonzero vectors then the rank of a matrix of the matrix $\boldsymbol{A} = \boldsymbol{u}\boldsymbol{v}^T$ is one. The range of $\boldsymbol{A}$ is the set of all vectors of the form

$$\boldsymbol{y} = \boldsymbol{A}\boldsymbol{x} = \boldsymbol{u}\boldsymbol{v}^T\boldsymbol{x} = (\boldsymbol{v}^T\boldsymbol{x})\boldsymbol{u}$$

since $\boldsymbol{u}$ is a nonzero vector, and not all vectors $\boldsymbol{v}^T\boldsymbol{x}$ are zero (because $\boldsymbol{v} \neq \boldsymbol{0}$) then this space is of dimension 1.

$\rightarrow$ Next we show that: If $\boldsymbol{A}$ is of rank one than there exist nonzero vectors $\boldsymbol{u}, \boldsymbol{v}$ such that $\boldsymbol{A} = \boldsymbol{u}\boldsymbol{v}^T$. If $\boldsymbol{A}$ is of rank one, then $\boldsymbol{Ran}(A) = \boldsymbol{Span}\{u\}$ for some nonzero vector $\boldsymbol{u}$. So for every vector $\boldsymbol{x}$, the vector $\boldsymbol{A}\boldsymbol{x}$ is a multiple of $\boldsymbol{u}$. Let $\boldsymbol{e}_1, \boldsymbol{e}_2, \cdots, \boldsymbol{e}_n$ the vectors of the canonical basis of $\mathbb{R}^n$ and let $\boldsymbol{\nu}_1, \boldsymbol{\nu}_2, \cdots, \boldsymbol{\nu}_n$ the scalars such that $\boldsymbol{A}\boldsymbol{e}_i = \boldsymbol{\nu}_i\boldsymbol{u}$. Define $\boldsymbol{v} = [\boldsymbol{\nu}_1, \boldsymbol{\nu}_2, \cdots, \boldsymbol{\nu}_n]^T$. Then $\boldsymbol{A} = \boldsymbol{u}\boldsymbol{v}^T$ because the matrices $\boldsymbol{A}$ and $\boldsymbol{u}\boldsymbol{v}^T$ have the same columns. (Note that the $\boldsymbol{j}$-th column of $\boldsymbol{A}$ is the vector $\boldsymbol{A}\boldsymbol{e}_j$). In addition, $\boldsymbol{v} \neq \boldsymbol{0}$ otherwise

$\boldsymbol{A} == \mathbf{0}$ which would be a contradiction because rank($\boldsymbol{A}$) = 1.

(b: second part) Eigenvalues /vectors

Write $\boldsymbol{Ax} = \boldsymbol{\lambda x}$ then notice that this means $(\boldsymbol{v^T x})\boldsymbol{u} = \boldsymbol{\lambda x}$ so either $\boldsymbol{v^T x} = \mathbf{0}$ and $\boldsymbol{\lambda} = \mathbf{0}$ or $\boldsymbol{x} = \boldsymbol{u}$ and $\boldsymbol{\lambda} = \boldsymbol{v^T u}$. Two eigenvalues: 0 and $\boldsymbol{v^T x}$... $\square$

✍17 Is it true that

$$\mathbf{rank}(\boldsymbol{A}) = \mathbf{rank}(\bar{\boldsymbol{A}}) = \mathbf{rank}(\boldsymbol{A^T}) = \mathbf{rank}(\boldsymbol{A^H}) \text{ ?}$$

Solution: The answer is yes and it follows from the fact that the ranks of $\boldsymbol{A}$ and $\boldsymbol{A^T}$ are the same and the ranks of $\boldsymbol{A}$ and $\bar{\boldsymbol{A}}$ are also the same.

It is known that $\boldsymbol{rank}(\boldsymbol{A}) = \boldsymbol{rank}(\boldsymbol{A^T})$. We now compare the ranks of $\boldsymbol{A}$ and $\bar{\boldsymbol{A}}$ (everything is considered to be complex).

The important property that is used is that if a set of vectors is linearly independent then so is its conjugate. [convince yourself of this by looking at material from 2033]. If $\boldsymbol{A}$ has rank $\boldsymbol{r}$ and for example its first $\boldsymbol{r}$ columns are the basis of the range, the the same $\boldsymbol{r}$ columns of $\bar{\boldsymbol{A}}$ are also linearly independent. So $\boldsymbol{rank}(\bar{\boldsymbol{A}}) \geq \boldsymbol{rank}(\boldsymbol{A})$. Now you can use a similar argument to show that $\boldsymbol{rank}(\boldsymbol{A}) \geq \boldsymbol{rank}(\bar{\boldsymbol{A}})$.

Therefore the ranks are the same. $\square$

✍21 Eigenvalues of two similar matrices A and B are the same. What about eigenvectors?

Solution: If $Au = \lambda u$ then $XBX^{-1}u = \lambda u \rightarrow B(X^{-1}u) = \lambda(X^{-1}u) \rightarrow \lambda$ is an eigenvalue of B with eigenvector $X^{-1}u$ (note that the vector $X^{-1}u$ cannot be equal to zero because $u \neq 0$.) $\square$

✍22 Given a polynomial $p(t)$ how would you define $p(A)$?

Solution: If $p(t) = \alpha_0 + \alpha_1 t + \alpha_2 t^2 + \cdots + \alpha_k t^k$ then

$$p(A) = \alpha_0 I + \alpha_1 A + \alpha_2 A^2 + \cdots + \alpha_k A^k \quad \text{where:}$$

$$A^j = \underbrace{A \times A \times \cdots \times A}_{j\text{times}}$$

$\square$

✍23 Given a function $f(t)$ (e.g., e^t) how would you define $f(A)$? [You may limit yourself to the case when A is diagonalizable]

Solution: The easiest way would be through the Taylor series expansion..

$$f(A) = f(0)I + \frac{f'(0)}{1!}A + \frac{f''(0)}{2!}A^2 \cdots \frac{f^{(k)}(0)}{k!}A^k + \cdots$$

However, this will require a justification: Will this expression 'converge' as the number of terms goes to infinity? This is where norms are useful. We will revisit this in next set. □

✍24 If A is nonsingular what are the eigenvalues/eigenvectors of A^{-1}?

Solution: Assume that $Au = \lambda u$. Multiply both sides by the inverse of A: $u = \lambda A^{-1}u$ - then by the inverse of λ: $\lambda^{-1}u = A^{-1}u$. Therefore, $1/\lambda$ is an eigenvalue and u is an associated eigenvector. □

✍25 What are the eigenvalues/eigenvectors of A^k for a given integer power k?

Solution: Assume that $Au = \lambda u$. Multiply both sides by A and repeat k times. You will get $A^k u = \lambda^k u$. Therefore, λ^k is an eigenvalue of A^k and u is an associated eigenvector. □

✍26 What are the eigenvalues/eigenvectors of $p(A)$ for a polynomial p?

Solution: Using the previous result you can show that $p(\lambda)$ is an eigenvalue of $p(A)$ and u is an associated eigenvector. $\square$

✍27 What are the eigenvalues/eigenvectors of $f(A)$ for a function f? [Diagonalizable case]

Solution: This will require using the diagonalized form of A: $A = XDX^{-1}$. With this $f(A) = Xf(D)X^{-1}$. It becomes clear that the eigenvalues are the diagonal entries of $f(D)$, i.e., the values $f(\lambda_i)$ for $i = 1, \cdots, n$. As for the eigenvectors - recall that they are the columns of the X matrix in the diagonalized form – And X is the same for A and $f(A)$. So the eigenvectors are the same. $\square$

✍28 For two $n \times n$ matrices A and B are the eigenvalues of AB and BA the same?

Solution: We will show that if λ is an eigenvalue of AB then it is also an eigenvalue of BA. Assume that $ABu = \lambda u$ and multiply both sides by B. Then $BABu = \lambda Bu$ – which we write in the form: $BAv = \lambda v$ where $v = Bu$. In the situation when $v \neq 0$, we clearly see that λ is a nonzero eigenvalue of BA with the associated eigenvector v. We now deal with the case when $v = 0$. In this case, since $ABu = \lambda u$, and $u \neq 0$ we must have $\lambda = 0$. However, clearly $\lambda = 0$ is also an eigenvalue of BA because $\det(BA) = det(AB) = 0$.

We can similarly show that any eigenvalue of BA are also eigenvalues of AB by interchanging the roles of A and B. This completes the proof ☐

✍29 Trace, spectral radius, and determinant of $A = \begin{pmatrix} 2 & 1 \\ 3 & 0 \end{pmatrix}$.

Solution: Trace is 2, determinant is -3. Eigenvalues are $3, -1$ so $\rho(A) = 3$. ☐

Exercises on determinants – all are rather straightforward except the following two.

✍35 Let A be a nonsingular diagonal $n \times n$ matrix. Show that:

$$\log \det(A) = \text{Trace}(\log(A))$$

Solution: If $A = \text{diag}(d_1, d_2, \cdots, d_n)$ then

$$\det(A) = \prod_{i=1}^{n} di_i \quad \rightarrow \quad \log \det(A) = \log \prod_{i=1}^{n} d_i = \sum_{i=1}^{n} \log d_i.$$

Now $f(A)$ is just the matrix with diagonal entrues $f(d_i)$.

$$\log \det(A) = \sum_{i=1}^{n} \log(d_i) = \text{Tr}\,(\log(A)).$$

☐

✍36 Let $C = \{c_{ij}\}_{i,j=1:n} \equiv$ matrix of cofactors. Show that $AC^T = \det(A) \times I$. So $A^{-1} =?$

Solution: Consider $(AC^T)_{ii}$ this is just the expansopn of det(A) w.r.t. row i shown in Page 1-38.

Consider $(AC^T)_{ij} = \sum a_{ik}c_{jk}$ with $j \neq i$. This is the expression of det(B) where B is obtained from A by copying the i=th row of A over its j-th row. This determinant is zero (since B has two identical rows). In the end $AC^T = \det(A) \times I$

This gives the well-known formula:

$$A^{-1} = \frac{1}{\det(A)} C^T.$$

☐

Basics on matrices [Csci2033 notes]

➤ If A is an $m \times n$ matrix (m rows and n columns) –then the scalar entry in the ith row and jth column of A is denoted by a_{ij} and is called the (i, j)-entry of A.

$$\begin{array}{ccc} & \text{Column } j & \\ & \downarrow & \\ \text{Row } i \rightarrow & \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots a_{1n} \\ \vdots & & \vdots & \vdots \\ a_{i1} & \cdots & \boxed{a_{ij}} & \cdots a_{in} \\ \vdots & & \vdots & \vdots \\ a_{m1} & \cdots & a_{mj} & \cdots a_{mn} \end{bmatrix} & = A \\ & \uparrow \quad\quad \uparrow \quad\quad \uparrow & \\ & a_{:1} \quad\quad a_{:j} \quad\quad a_{:n} & \end{array}$$

➤ a_{ij} == ith entry (from the top) of the jth column

➤ Each column of A is a list of m real numbers, which identifies a vector in $\mathbb{R}^m$ called a column vector

➤ The columns $a_{:1}..., a_{:n}$ - denoted by $a_1, a_2, \cdots, a_n$ so $A = [a_1, a_2, \cdots, a_n]$

➤ The diagonal entries in an $m \times n$ matrix A are a_{11}, a_{22}, a_{33}. They form the main diagonal of A.

➤ A diagonal matrix is a matrix whose nondiagonal entries are zero

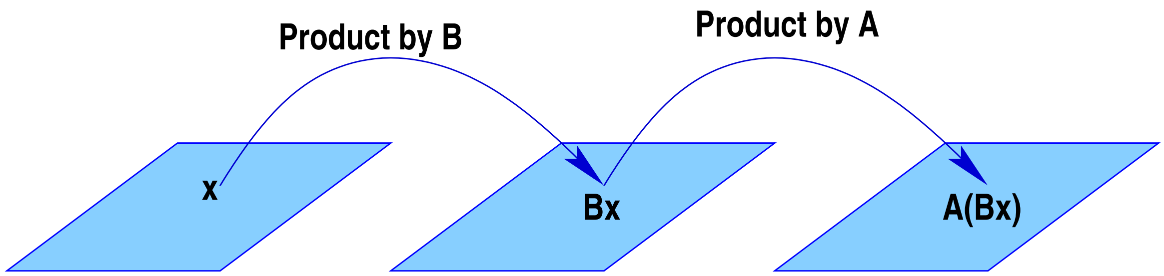
➤ The $n \times n$ identity matrix I_n Example:

$$I_3 = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

Matrix Multiplication

➤ When a matrix B multiplies a vector x, it transforms x into the vector Bx.

➤ If this vector is then multiplied in turn by a matrix A, the resulting vector is $A(Bx)$.



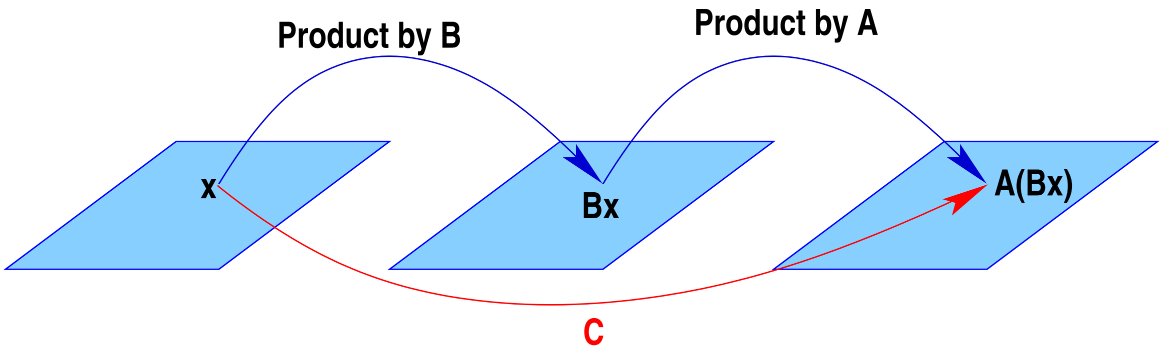
➤ Thus $A(Bx)$ is produced from x by a composition of mappings–the linear transformations induced by B and A.

➤ Note: $x \rightarrow A(Bx)$ is a linear mapping (prove this).

Goal: to represent this composite mapping as a multiplication by a single matrix, call it C for now, so that

$$A(Bx) = Cx$$

.



➤ Assume A is $m \times n$, B is $n \times p$, and x is in $\mathbb{R}^p$. Denote the columns of B by $b_1, \cdots, b_p$ and the entries in x by $x_1, \cdots, x_p$. Then:

$$Bx = x_1 b_1 + \cdots + x_p b_p$$

➤ By the linearity of multiplication by A:

$$\begin{aligned} A(Bx) &= A(x_1 b_1) + \cdots + A(x_p b_p) \\ &= x_1 A b_1 + \cdots + x_p A b_p \end{aligned}$$

➤ The vector $A(Bx)$ is a linear combination of Ab_1, $\cdots$, Ab_p, using the entries in x as weights.

➤ Matrix notation: this linear combination is written as

$$A(Bx) = [Ab_1, Ab_2, \cdots Ab_p].x$$

➤ Thus, multiplication by $[Ab_1, Ab_2, \cdots, Ab_p]$ transforms x into $A(Bx)$.

➤ Therefore the desired matrix C is the matrix

$$C = [Ab_1, Ab_2, \cdots, Ab_p]$$

Definition: If A is an $m \times n$ matrix, and if B is an $n \times p$ matrix with columns $b_1, \cdots, b_p$, then the product AB is the matrix whose p columns are $Ab_1, \cdots, Ab_p$. That is:

$$AB = A[b_1, b_2, \cdots, b_p] = [Ab_1, Ab_2, \cdots, Ab_p]$$

➤ Remeber

Multiplication of matrices corresponds to composition of linear transformations.

✍ Operation count: How many operations are required to perform product AB?

✍ Compute AB when:

$$A = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 2 & -1 \\ 1 & 3 & -2 \end{bmatrix}$$

✍ Compute AB when:

$$A = \begin{bmatrix} 2 & -1 & 2 & 0 \\ 1 & -2 & 1 & 0 \\ 3 & -2 & 0 & 0 \end{bmatrix} \quad B = \begin{bmatrix} 1 & -1 & -2 \\ 0 & -2 & 2 \\ 2 & 1 & -2 \\ -1 & 3 & 2 \end{bmatrix}$$

✍ Can you compute AB when:

$$A = \begin{bmatrix} 2 & -1 \\ 1 & 3 \end{bmatrix} \quad B = \begin{bmatrix} 0 & 2 \\ 1 & 3 \\ -1 & 4 \end{bmatrix} ?$$

Row-wise matrix product

➤ Recall what we did with matrix-vector product to compute a single entry of the vector Ax

➤ Can we do the same thing here? i.e., How can we compute the entry c_{ij} of the product AB without computing entire columns?

✍ Do this to compute entry $(2, 2)$ in the first example above.

✍ Operation counts: Is more or less expensive to perform the matrix-vector product row-wise instead of column-wise?

1-16

Properties of matrix multiplication

Theorem Let A be an $m \times n$ matrix, and let B and C have sizes for which the indicated sums and products are defined. Then:

- $A(BC) = (AB)C$ (associative law of multiplication)
- $A(B + C) = AB + AC$ (left distributive law)
- $(B + C)A = BA + CA$ (right distributive law)
- $\alpha(AB) = (\alpha A)B = A(\alpha B)$ for any scalar α
- $I_m A = AI_n = A$ (product with identity)

✍ If $AB = AC$ then $B = C$ ('simplification') : True-False?

✍ If $AB = 0$ then either $A = 0$ or $B = 0$: True or False?

✍ $AB = BA$: True or false??

Square matrices. Matrix powers

➤ Important particular case when $n = m$ - so matrix is $n \times n$

➤ In this case if x is in $\mathbb{R}^n$ then $y = Ax$ is also in $\mathbb{R}^n$

➤ AA is also a square $n \times n$ matrix and will be denoted by A^2

➤ More generally, the matrix A^k is the matrix which is the product of k copies of A:

$$A^1 = A; \quad A^2 = AA; \quad \cdots \quad A^k = \underbrace{A \cdots A}_{k \text{ times}}$$

➤ For consistency define A^0 to be the identity: $A^0 = I_n$,

✍ $A^l \times A^k = A^{l+k}$ – Also true when k or l is zero.

Transpose of a matrix

Given an $m \times n$ matrix A, the transpose of A is the $n \times m$ matrix, denoted by A^T, whose columns are formed from the corresponding rows of A.

Theorem : Let A and B denote matrices whose sizes are appropriate for the following sums and products. Then: • $(A^T)^T = A$
• $(A + B)^T = A^T + B^T$
• $(\alpha A)^T = \alpha A^T$ for any scalar α
• $(AB)^T = B^T A^T$

More on matrix products

➤ Recall: Product of the matrix A by the vector x: (a_j == jth column of A)

$$\underset{y}{\begin{bmatrix}\beta_1\\ \vdots\\ \beta_j\\ \vdots\\ \beta_n\end{bmatrix}} = \underset{A}{\begin{bmatrix}a_{11} & \cdots & a_{1j} & \cdots a_{1n}\\ \vdots & & \vdots & \vdots\\ a_{i1} & \cdots & a_{ij} & \cdots a_{in}\\ \vdots & & \vdots & \vdots\\ a_{m1} & \cdots & a_{mj} & \cdots a_{mn}\end{bmatrix}} \underset{x}{\begin{bmatrix}\alpha_1\\ \vdots\\ \alpha_j\\ \vdots\\ \alpha_n\end{bmatrix}}$$

$$= \alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n$$

- x, y are vectors; y is the result of $A \times x$.
- $a_1, a_2, ..., a_n$ are the columns of A
- $\alpha_1, \alpha_2, ..., \alpha_n$ are the components of x [scalars]
- $\alpha_1 a_1$ is the first column of A multiplied by the scalar α_1 which is the first component of x.
- $\alpha_1 a_1 + \alpha_2 a_2 + \cdots + \alpha_n a_n$ is a linear combination of $a_1, a_2, \cdots, a_n$ with weights $\alpha_1, \alpha_2, ..., \alpha_n$.

➤ This is the 'column-wise' form of the 'matvec'

Example:

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix} \quad x = \begin{bmatrix} -2 \\ 1 \\ -3 \end{bmatrix} \quad y = ?$$

➤ Result:

$$y = -2 \times \begin{bmatrix} 1 \\ 0 \end{bmatrix} + 1 \times \begin{bmatrix} 2 \\ -1 \end{bmatrix} - 3 \times \begin{bmatrix} -1 \\ 3 \end{bmatrix} = \begin{bmatrix} 3 \\ -10 \end{bmatrix}$$

➤ Can get i-th component of the result y without the others: $\beta_i = \alpha_1 a_{i1} + \alpha_2 a_{i2} + \cdots + \alpha_n a_{in}$

Example: In the above example extract β_2

$$\beta_2 = (-2) \times 0 + (1) \times (-1) + (-3) \times (3) = -10$$

➤ Can compute $\beta_1, \beta_2, \cdots, \beta_m$ in this way.

➤ This is the 'row-wise' form of the 'matvec'

Matrix-Matrix product

➤ Recall:

➤ When A is $m \times n$, B is $n \times p$, the product AB of the matrices A and B is the $m \times p$ matrix defined as

$$AB = [Ab_1, Ab_2, \cdots, Ab_p]$$

where $b_1, b_2, \cdots, b_p$ are the columns of B

➤ Each Ab_j == product of A by the j-th column of B. Matrix AB is in $\mathbb{R}^{m \times p}$

➤ Can use what we know on matvecs to perform the product

1. Column form – In words: *"The j-th column of AB is a linear combination of the columns of A, with weights $b_{1j}, b_{2j}, \cdots, b_{nj}$ "* (entries of j-th col. of B)

Example:

$$A = \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix} \quad B = \begin{bmatrix} -2 & 1 \\ 1 & -2 \\ -3 & 2 \end{bmatrix} \quad AB = ?$$

➤ Result:

$$B = \left[\begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix} \begin{bmatrix} -2 \\ 1 \\ -3 \end{bmatrix}, \begin{bmatrix} 1 & 2 & -1 \\ 0 & -1 & 3 \end{bmatrix} \begin{bmatrix} 1 \\ -2 \\ 2 \end{bmatrix} \right]$$

$$= \begin{bmatrix} 3 & -6 \\ -10 & 8 \end{bmatrix}$$

➤ First column has been computed before: it is equal to:
(-2)*(col. 1 of A) + (1)*(col. 2 of A) + (-3)*(col. 3 of A)

➤ Second column is equal to:
(1)*(col. 1 of A) + (-2)*(col. 2 of A) + (2)*(col. 3 of A)

2. If we call C the matrix $C = AB$ what is c_{ij}? From above:

$$c_{ij} = a_{i1}b_{1j} + a_{i2}b_{2j} + \cdots + a_{ik}b_{kj} + \cdots + a_{in}b_{nj}$$

➤ Fix j and run $i \longrightarrow$ column-wise form just seen

3. Fix i and run $j \longrightarrow$ row-wise form

Example: Get second row of AB in previous example.

$$c_{2j} = a_{21}b_{1j} + a_{22}b_{2j} + a_{23}b_{3j}, \quad j = 1, 2$$

- Can be read as : $\boxed{c_{2:} = a_{21}b_{1:} + a_{22}b_{2:} + a_{23}b_{3:}}$, or in words:

row2 of C = a_{21} (row1 of B) + a_{22} (row2 of B) + a_{23} (row3 of B)

= 0 (row1 of B) + (-1) (row2 of B) + (3) (row3 of B)

= $[-10 \quad 8]$