Review Packet 1

1. For each of the following, write the vector or matrix that is specified:

   a. $\mathbf{e}_3 \in \mathbb{R}^4$

   b. $\mathbf{D} = diag\{2, \sqrt{3}, -1\}$

   c. $\mathbf{e} \in \mathbb{R}^3$

   d. $\mathbf{I}_2$

a) $\begin{pmatrix} 0 \\ 0 \\ 1 \\ 0 \end{pmatrix}$ b) $\begin{pmatrix} 2 & 0 & 0 \\ 0 & \sqrt{3} & 0 \\ 0 & 0 & -1 \end{pmatrix}$ c) $\begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix}$ d) $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

2. For each of the following matrices and vectors, give their dimension. Label each as a matrix or vector. For each matrix, indicate whether the matrix is square or rectangular.

   a.
$$\mathbf{A} = \begin{pmatrix} 2 & 3 & -1 \\ 1 & -1 & 1 \\ 2 & 2 & 1 \end{pmatrix}$$
3x3 square matrix

   b.
$$\mathbf{h} = \begin{pmatrix} -1 \\ -4 \\ 1 \\ 2 \end{pmatrix}$$
4x1 vector

   c.
$$\mathbf{B} = \begin{pmatrix} B_{11} & B_{12} & B_{13} \\ B_{21} & B_{22} & B_{23} \\ B_{31} & B_{32} & B_{33} \\ B_{41} & B_{42} & B_{43} \end{pmatrix}$$
4x3 rectangular matrix

   d.
$$\mathbf{A} = [A_{ij}] \quad \text{where } i = 1,2,3 \quad \text{and } j = 1,2$$
3x2 rectangular matrix

3. Specify whether the following augmented matrices are in row-echelon form (REF), reduced row-echelon form (RREF), or neither:

   a. $\left(\begin{array}{ccc|c} 3 & 2 & 1 & 2 \\ 0 & 2 & 0 & 1 \\ 0 & 0 & 1 & 5 \end{array}\right)$ REF (could eliminate (1,2)-element and (1,3)-element)

   b. $\left(\begin{array}{ccc|c} 3 & 2 & 1 & 2 \\ 0 & 2 & 0 & 1 \\ 0 & 4 & 0 & 0 \end{array}\right)$ neither

   c. $\left(\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$ RREF

   d. $\left(\begin{array}{ccc|c} 1 & 2 & 0 & 2 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$ REF

4. *(True/False)* The normal equations are used to find the ordinary least-squares solution to an inconsistent system of equations.

TRUE.

5. If the matrix equation $\mathbf{Mv} = \mathbf{b}$ is inconsistent, what alternative equation should I solve to find a solution $\widehat{\mathbf{v}}$ such that $\mathbf{M}\widehat{\mathbf{v}} = \widehat{\mathbf{b}}$ is as close to $\mathbf{b}$ as possible in the sense that it minimizes the sum of squared error:

$$SSE = \sum_{i=1}^{n} (\widehat{\mathbf{b}}_i - \mathbf{b}_i)^2$$

$M^T M \hat{v} = M^T b$ (the normal eqns)

6. Answer the following questions about each matrix

$$\mathbf{A} = \begin{pmatrix} 4 & 5 & 2 \\ 5 & 3 & 1 \\ 2 & 1 & -2 \end{pmatrix} \qquad \mathbf{B} = \begin{pmatrix} 1 & 0 & 1 & -1 \\ 0 & 2 & 1 & 0 \\ 0 & 1 & -1 & 1 \end{pmatrix}$$

a. Is the matrix square?

A Yes **B** NO

b. What is the transpose of the matrix?

$\mathbf{A}^T = A$ $\mathbf{B}^T = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 2 & 1 \\ 1 & 1 & -1 \\ -1 & 0 & 1 \end{pmatrix}$

c. Is the matrix symmetric?

A yes **B** no

d. If possible, name the diagonal elements of the matrix.

A 4, 3, -2 **B** n.p.

e. If possible, compute the Trace of the matrix.

A $4+3-2 = 5$ **B** n.p.

f. Can the product **AB** be computed? If so, what is the size of the result?

yes 3x4

g. Can the product **BA** be computed? If so, what is the size of the result?

NO

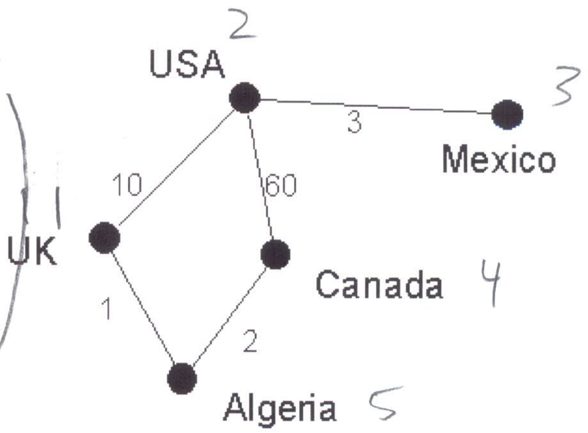
h. Can the product $(\mathbf{B}_{\star 3})^T(\mathbf{A}_{3\star})^T$ be computed? If so, what is the result? Yes $(1\ 1\ -1)\begin{pmatrix} 2 \\ 1 \\ -2 \end{pmatrix} = 5$

7. What is the inverse of the matrix $\mathbf{D} = \sigma \mathbf{I}_3$?

$$D^{-1} = \frac{1}{\sigma} I_3 = \begin{pmatrix} 1/\sigma & 0 & 0 \\ 0 & 1/\sigma & 0 \\ 0 & 0 & 1/\sigma \end{pmatrix}$$

8. For the following graph, number the nodes and write the corresponding adjacency matrix:

$$\begin{array}{c|c} & \begin{array}{ccccc} \overset{1}{\text{UK}} & \overset{2}{\text{USA}} & \overset{3}{\text{MEX}} & \overset{4}{\text{CAN}} & \overset{5}{\text{ALG}} \end{array} \\ \begin{array}{c} \text{K} \\ \text{A} \\ \text{EX} \\ \text{N} \\ \text{G} \end{array} & \begin{pmatrix} 0 & 10 & 0 & 0 & 1 \\ 10 & 0 & 3 & 60 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 60 & 0 & 0 & 2 \\ 1 & 0 & 0 & 2 & 0 \end{pmatrix} \end{array}$$



ANSWER DEPENDS ON ORDER OF NODES

9. Compute the outer product $\mathbf{x}\mathbf{y}^T$ where

$$\mathbf{x} = \begin{pmatrix} 3 \\ 4 \\ 5 \end{pmatrix} \quad \mathbf{y} = \begin{pmatrix} 1 \\ 1 \\ -1 \\ 1 \end{pmatrix}$$

$$\mathbf{x}\mathbf{y}^T = \begin{pmatrix} 3 & 3 & -3 & 3 \\ 4 & 4 & -4 & 4 \\ 5 & 5 & -5 & 5 \end{pmatrix}$$

10. Can you say anything about the rank of an outer product in general? Explain your answer.

It will always equal 1 b/c the rows will be multiples of each other ⇒ one linearly ind. row.

11. Briefly explain what it means for a matrix to be full rank.

Either the rows are linearly independent or the columns are linearly independent

$A_{m \times n}$ then $\text{rank}(A) = \min(m,n)$

12. For the following augmented matrices, circle the pivot elements and give the rank of the coefficient matrix along with the number of free variables.

a. $\left(\begin{array}{cccc|c} 3 & 2 & 1 & 1 & 2 \\ 0 & 2 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 5 \end{array}\right)$ rank= 3 # free var= 1

b. $\left(\begin{array}{ccc|c} 1 & 1 & 0 & 2 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array}\right)$ rank= 2 # free var= 1

c. $\left(\begin{array}{ccccc|c} 1 & 2 & 0 & 1 & 0 & 2 \\ 0 & 1 & 1 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{array}\right)$ rank= 3 # free var= 2

13. Write the vector $\mathbf{v}$ as a linear combination of each given $\mathbf{x}$ and $\mathbf{y}$, if possible.

$$\mathbf{v} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$$

a. $\mathbf{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}$ $\mathbf{y} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}$ $V = 2x + 3y$

b. $\mathbf{x} = \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ $\mathbf{y} = \begin{pmatrix} 0 \\ -1 \end{pmatrix}$ $V = -2x - 3y$

c. $\mathbf{x} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ $\mathbf{y} = \begin{pmatrix} 2 \\ 2 \end{pmatrix}$ not possible

$V = ax + by \Leftrightarrow \begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix}\begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} 2 \\ 3 \end{pmatrix}$ system is in consistent

14. Suppose we measure the heights of 10 people, $person_1, person_2, \ldots, person_{10}$.

a. If we create a matrix $\mathbf{S}$ where

$$S_{ij} = height(person_i) - height(person_j)$$

is the matrix $\mathbf{S}$ symmetric? What is the trace($\mathbf{S}$)?

NO. $S_{ij} = -S_{ji}$ so not symmetric. $\text{Trace}(S) = 0$ since $S_{ii} = 0$ for all i.

b. If instead we create a matrix $\mathbf{G}$ where

$$G_{ij} = [height(person_i) - height(person_j)]^2$$

is the matrix $\mathbf{G}$ symmetric? What is the trace($\mathbf{G}$)?

Yes. $G_{ij} = G_{ji} \Rightarrow$ symmetric. $\text{Trace}(G) = 0$

15. For the matrix $\mathbf{H} = \mathbf{X}(\mathbf{X}^T\mathbf{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{X}^T\mathbf{\Sigma}^{-1}$, use the properties of matrix arithmetic to show that

a. $\mathbf{H}^2 = \mathbf{H}$

$H^2 = HH = X(X^T\Sigma^{-1}X)^{-1}\cancel{X^T\Sigma^{-1}X(X^T\Sigma^{-1}X)^{-1}}X^T\Sigma^{-1}$

$= X(X^T\Sigma^{-1}X)^{-1}X^T\Sigma^{-1} = H$ ✓

b. $\mathbf{H}(\mathbf{I} - \mathbf{H}) = \mathbf{0}$

$= H - H^2 = 0$
by part a!

16. Let

$$\mathbf{U} = (\mathbf{U}_1|\mathbf{U}_2|\mathbf{U}_3|\dots|\mathbf{U}_p) \quad \text{and} \quad \mathbf{V}^T = \begin{pmatrix} \mathbf{V}_1^T \\ \hline \mathbf{V}_2^T \\ \hline \mathbf{V}_3^T \\ \hline \vdots \\ \hline \mathbf{V}_p^T \end{pmatrix}$$

Write the matrix product $\mathbf{U}\mathbf{V}^T$ in terms of the columns of $\mathbf{U}$ and the rows of $\mathbf{V}^T$.

$UV^T = U_1V_1^T + U_2V_2^T + U_3V_3^T + \dots + U_pV_p^T$

17. Suppose that $\mathbf{u}$ is a unit vector. Then, $\|\mathbf{u}\|_2 =$?

1. That's the def. of a unit vector!

18. Let $\mathbf{x} = \begin{pmatrix} 1 \\ -1 \\ 2 \end{pmatrix}$ and $\mathbf{y} = \begin{pmatrix} 0 \\ -1 \\ -3 \end{pmatrix}$. Compute the following:

a. $\|\mathbf{x}\|_2$ $\sqrt{1^2+(-1)^2+(2)^2} = \sqrt{6}$

b. $\|\mathbf{y}\|_1$ $= |0| + |-1| + |-3| = 4$

c. $\|\mathbf{y}\|_\infty$ $\max\{|0|, |-1|, |-3|\} = 3$

d. $\|\mathbf{x}-\mathbf{y}\|_2$ $\sqrt{(1-0)^2 + (-1+1)^2 + (2+3)^2} = \sqrt{26}$

e. $\|\mathbf{x}-\mathbf{y}\|_1$ $|1-0| + |-1+1| + |2+3| = 6$

19. Suppose we have a dataset containing survey data. Individuals were asked to respond 'yes'=1 or 'no'=0 to twenty potential political referendums. Let $\mathbf{a}$ be the vector containing the numerical responses of Individual A and let $\mathbf{b}$ be the vector containing the numerical responses of Individual B (so $\mathbf{a}, \mathbf{b} \in \mathbb{R}^{20}$). Explain in words the interpretation of the quantity

$$\|\mathbf{a} - \mathbf{b}\|_1.$$

the number of referendums upon which individ. A and Indiv. B did not agree.

20. **Statistical Formulas Using Linear Algebra Notation.** Almost every statistical formula can be written in a more compact fashion using linear algebra. Most of the elementary formulas involve vector inner products or the Euclidean norm. To begin, we'll introduce the concept of *centering* the data. **Centering** the data means that the mean of a variable is subtracted from each observation. For example, if we have some variable, $\mathbf{x}$, and 3 observations on that variable:

$$\mathbf{x} = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix}$$

then obviously, $\bar{x} = 3$. The **centered** version of $\mathbf{x}$ would then be

$$\mathbf{x} - \bar{x}\mathbf{e} = \begin{pmatrix} 2 \\ 3 \\ 4 \end{pmatrix} - \begin{pmatrix} 3 \\ 3 \\ 3 \end{pmatrix} = \begin{pmatrix} -1 \\ 0 \\ 1 \end{pmatrix}.$$

We simply subtract the mean from every observation so that the new mean of the variable is 0.

Most multivariate textbooks start by saying *"all variable vectors in this textbook are assumed to be centered to have mean zero unless otherwise specified"*. Looking at the most common statistical formulas helps us see why. Try to re-write the following formulas using linear algebra notation, using the vectors $\mathbf{x}$ and $\mathbf{y}$ to represent centered data:

$$\mathbf{x} = \begin{pmatrix} x_1 - \bar{x} \\ x_2 - \bar{x} \\ x_3 - \bar{x} \\ \vdots \\ x_n - \bar{x} \end{pmatrix}, \quad \mathbf{y} = \begin{pmatrix} y_1 - \bar{y} \\ y_2 - \bar{y} \\ y_3 - \bar{y} \\ \vdots \\ y_n - \bar{y} \end{pmatrix}$$

For this exercise, keep in mind the following linear algebra constructs, which you should be very familiar with by now:

$$\|\mathbf{a}\| = \sqrt{a_1^2 + a_2^2 + a_3^2 + \cdots + a_n^2}$$

$$\mathbf{a}^T\mathbf{b} = a_1b_1 + a_2b_2 + a_3b_3 + \cdots + a_nb_n$$

a. Sample standard deviation:

$$s = \frac{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2}}{\sqrt{n-1}} = \boxed{\frac{\|x\|}{\sqrt{n-1}}}$$

b. Sample covariance:

$$covariance(\mathbf{x}, \mathbf{y}) = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y}) = \boxed{\frac{1}{n-1} x^T y}$$

c. Correlation coefficient:

$$r_{xy} = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}} = \boxed{\frac{x^T y}{\|x\| \cdot \|y\|}}$$

note: this is just the inner product of unit vectors!

$$\left(\frac{x}{\|x\|}\right)^T \left(\frac{y}{\|y\|}\right)$$