Recitation 1. February 12

Focus: recognizing vector spaces, rules of matrix multiplication.

Definition. A *real vector space* V is a set endowed with operations of adding two vectors and multiplying a vector by a real number such that the following holds for any three vectors u, v, w and any real numbers a, b:

- $(u+v)+w = u+(v+w)$;
- $u+v = v+u$;
- there exists a *zero vector* $0 \in V$ such that $v+0 = v$;
- $a(bv) = (ab)v$;
- $1v = v$;
- $(a+b)v = av+bv$;
- $a(u+v) = au+av$.

Elements of a vector space are called *vectors*.

Remark. The most basic rule that you should remember: $\boxed{\textbf{row column}}$. It shows the order in which you write or compute, e.g.:

- The first index denotes the row, the second number the column.
- You multiply a row by a column to get a number.
- An $n \times m$ matrix has n rows and m columns.

Notation. We will denote by A^T the transpose of a matrix A.

1. Is this a vector space? Why / why not? Which natural operations you considered when checking axioms?

   a) The line $y = x$.
   b) The line $y = x+1$.
   c) The union of the x and y axes.
   d) The unit circle $\{(x, y) \mid x^2 + y^2 = 1\}$.
   e) The set of 5×5 matrices with the element in position $(3, 3)$ being 0.
   f) Functions of the form $f(x) = ax^2 + bx + c$.
   g) Functions $f(x)$ with $f(7) = 0$.
   h) Functions $f(x)$ with $f(0) = 7$.

2. *Rules of matrix multiplication. (Section 2.4 of Strang.)* Consider the following matrices:

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \end{pmatrix},\ B = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 3 & 2 \end{pmatrix},\ C = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix},\ D = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix},\ E = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

   Which of these matrix operations are allowed, and what are the results?

   a) AB
   b) AB^T
   c) B^TA
   d) $(A+B)C$
   e) $(A+B)C^T$
   f) $C(A+B)$
   g) DB

h) BD

i) AE

j) EA

k) CAE

3. When you multiply an $n \times m$ matrix by an $m \times l$ matrix, what are the dimensions of the resulting matrix?

Solution:

4. When can a matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be written as $X^T X$ for some other matrix $X = \begin{pmatrix} p & 0 \\ q & r \end{pmatrix}$? Assume that $b \neq 0$. What are p, q, r in terms of a, b, c, d when possible?

Solution:

5. Consider the following matrices:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, B = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}, C = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, D = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}.$$

How is each row of BA, CA, DA related to the rows of A? How is each column of AB, AC, AD related to the columns of A?

Solution:

6. *In this problem, we will practice block multiplication. (Page 75 of Strang.)* Consider the following column vector c and a 3×3 matrix A with columns a_1, a_2, a_3:

$$c = \begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix}, A = \begin{pmatrix} | & | & | \\ a_1 & a_2 & a_3 \\ | & | & | \end{pmatrix}.$$

Write the result of matrix multiplication rA as a linear combination of the column vectors a_1, a_2, a_3. What if we write a matrix R as three rows $R = \begin{pmatrix} - & r_1 & - \\ - & r_2 & - \\ - & r_3 & - \end{pmatrix}$ and multiply R by A?

Solution:

Recitation 1. Solution

Focus: recognizing vector spaces, rules of matrix multiplication.

We will provide a formal definition of a vector space for the sake of honesty, but for the purpose of solving the first problem of this worksheet, you will need to check just three properties: when you add two elements, you get an element of the same set (being closed under addition); when you multiply an element by a scalar, the result is also in the same set (being closed under multiplication by scalars); and zero should belong to this set. These properties are boxed. Once again: we do not expect you to be able to recite all the axioms, although we definitely would be impressed if you are :)

Definition. A *real vector space* V is a set endowed with operations of $\boxed{\text{adding two vectors}}$ and $\boxed{\text{multiplying a vector by a real number}}$ such that the following holds for any three vectors u, v, w and any real numbers a, b:

- $(u + v) + w = u + (v + w)$;
- $u + v = v + u$;
- $\boxed{\text{there exists a } \textit{zero vector } 0 \in V \text{ such that for any vector } v\text{, we have } v + 0 = v}$;
- $a(bv) = (ab)v$;
- $1v = v$;
- $(a + b)v = av + bv$;
- $a(u + v) = au + av$.

Elements of a vector space are called *vectors*.

Remark. The most basic rule that you should remember: $\boxed{\textbf{row column}}$. It shows the order in which you write or compute, e.g.:

- The first index denotes the row, the second number the column.
- You multiply a row by a column to get a number.
- An $n \times m$ matrix has n rows and m columns.

Notation. We will denote by A^T the transpose of a matrix A.

1. Is this a vector space? Why / why not? Which natural operations you considered when checking axioms?

   a) The line $y = x$.
   b) The line $y = x + 1$.
   c) The union of the x and y axes.
   d) The unit circle $\{(x, y) \mid x^2 + y^2 = 1\}$.
   e) The set of 5×5 matrices with the element in position $(3, 3)$ being 0.
   f) Functions of the form $f(x) = ax^2 + bx + c$.
   g) Functions $f(x)$ with $f(7) = 0$.
   h) Functions $f(x)$ with $f(0) = 7$.
   i) *Tricky question.* Newtonian universe.

Solution:

a) Yes.

b) No, because the set is not closed under addition. For example, the points $\binom{-1}{0}$ and $\binom{0}{1}$ belong to the line, but their sum $\binom{-1}{1}$ does not. Neither is the set closed under multiplication or has zero.

c) No, because the set is not closed under addition. Example of points is the same as above. However, it is closed under scaling and contains zero.

d) No, for the same reasons as in part (b).

e) Yes. Adding two matrices or multiplying such a matrix by a number does not affect the property that the middle element is zero.

f) Yes, the set of quadratic polynomials is a vector space.

g) Yes.

h) No, because the set is not closed under addition: if you add two functions f and g with $f(0) = g(0) = 7$, then their sum evaluates to 14 at 0.

i) No, we don't have a zero, because there is no natural reference point.

2. *Rules of matrix multiplication. (Section 2.4 of Strang.)* Consider the following matrices:

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \end{pmatrix},\ B = \begin{pmatrix} -1 & 0 & 1 \\ -1 & 3 & 2 \end{pmatrix},\ C = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix},\ D = \begin{pmatrix} 1 & 2 & 3 \end{pmatrix},\ E = \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix}.$$

Which of these matrix operations are allowed, and what are the results?

a) AB
b) AB^T
c) B^TA
d) $(A+B)C$
e) $(A+B)C^T$
f) $C(A+B)$
g) DB
h) BD
i) AE
j) EA
k) CAE

Solution: In order to multiply two matrices, number of columns in the first should be equal to number of rows in the second.

a) AB not allowed: we cannot multiply a 2×3 matrix by a 2×3 matrix.

b) $AB^T = \begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \end{pmatrix} \begin{pmatrix} -1 & -1 \\ 0 & 3 \\ 1 & 2 \end{pmatrix} = \begin{pmatrix} -1 & 5 \\ -1 & -2 \end{pmatrix}.$

c) $B^TA = \begin{pmatrix} -1 & -1 \\ 0 & 3 \\ 1 & 2 \end{pmatrix} \begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \end{pmatrix} = \begin{pmatrix} -4 & -1 & -2 \\ 9 & -3 & 6 \\ 7 & 0 & 4 \end{pmatrix}.$

d) $(A+B)C$ not allowed.

e) $(A+B)C^T$ not allowed.

f) $C(A+B) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} 0 & 2 & 1 \\ 2 & 2 & 4 \end{pmatrix} = \begin{pmatrix} 0 & -4 & -2 \\ -4 & -4 & -8 \end{pmatrix}$.

g) DB not allowed.

h) BD not allowed.

i) $AE = \begin{pmatrix} 1 & 2 & 0 \\ 3 & -1 & 2 \end{pmatrix} \begin{pmatrix} 3 \\ 2 \\ 1 \end{pmatrix} = \begin{pmatrix} 7 \\ 9 \end{pmatrix}$.

j) EA not allowed.

k) $CAE = C(AE) = \begin{pmatrix} -2 & 0 \\ 0 & -2 \end{pmatrix} \begin{pmatrix} 7 \\ 9 \end{pmatrix} = \begin{pmatrix} -14 \\ -18 \end{pmatrix}$.

3. When you multiply an $n \times m$ matrix by an $m \times l$ matrix, what are the dimensions of the resulting matrix?

Solution: $(n \times m)(m \times l) \to (n \times l)$.

4. When can a matrix $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ be written as $X^T X$ for some other matrix $X = \begin{pmatrix} p & 0 \\ q & r \end{pmatrix}$? Assume that $b \neq 0$. What are p, q, r in terms of a, b, c, d when possible?

Solution: In order to describe the condition, we compare matrix elements in the desired equality $A = X^T X$:

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} p & q \\ 0 & r \end{pmatrix} \begin{pmatrix} p & 0 \\ q & r \end{pmatrix} = \begin{pmatrix} p^2+q^2 & qr \\ qr & r^2 \end{pmatrix}.$$

One can note that matrix elements of $X^T X$ are dot products of the columns of X.

One can immediately observe that A should satisfy $a \geq 0$ and $d \geq 0$, because square of a real number is always nonnegative, and $b = qr = c$, that is A is necessarily symmetric. In addition, if either a or d is zero, then $b = c = 0$ as well, because vanishing of a implies that $q = 0$, so $b = c = qr = 0$, and similarly for d. There will be one more condition which we will find later.

Now we turn to expressing p, q and r in terms of a, b, c and d:

- $r = \sqrt{d}$;
- if $d \neq 0$, then $q = \frac{b}{r} = \frac{b}{\sqrt{d}}$ and $p = \sqrt{a - q^2} = \sqrt{a - \frac{b^2}{d}}$, hence we have an additional constraint that $a - \frac{b^2}{d} \geq 0$;
- if $d = 0$, then $r = 0$ and p and q are any real numbers such that the vector $\begin{pmatrix} p \\ q \end{pmatrix}$ lies on the circle of radius a.

5. Consider the following matrices:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, B = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}, C = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, D = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}.$$

How is each row of BA, CA, DA related to the rows of A? How is each column of AB, AC, AD related to the columns of A?

Solution: I will only write the solution for rows, because what happens to columns is exactly the same after you change the order of multiplication.

- The first row of BA is twice the first row of A, and the second is minus the second row of A.
- The first row of CA is the second row of A, while the second row is zero.
- The first row of DA is the second row of A and the second row of DA is minus the second row of A.

So you can see that multiplying a matrix A by another matrix on the left performs row operations. Similarly, right multiplication performs column operations. You will see more in the next problem.

6. *In this problem, we will practice block multiplication. (Page 75 of Strang.)* Consider the following column vector c and a 3×3 matrix A with columns a_1, a_2, a_3:

$$c = \begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix}, \; A = \begin{pmatrix} | & | & | \\ a_1 & a_2 & a_3 \\ | & | & | \end{pmatrix}.$$

Write the result of matrix multiplication rA as a linear combination of the column vectors a_1, a_2, a_3. What if we write a matrix R as three rows $R = \begin{pmatrix} - & r_1 & - \\ - & r_2 & - \\ - & r_3 & - \end{pmatrix}$ and multiply R by A?

Solution: First compute Ac, and note that a_1, a_2, a_3 are all 3-vectors:

$$Ac = \begin{pmatrix} | & | & | \\ a_1 & a_2 & a_3 \\ | & | & | \end{pmatrix} \begin{pmatrix} \lambda \\ \mu \\ \nu \end{pmatrix} = \lambda a_1 + \mu a_2 + \nu a_3.$$

This is a particular case of block multiplication.

Now calculate RA – this turns out to be the usual rule of matrix multiplication:

$$RA = \begin{pmatrix} - & r_1 & - \\ - & r_2 & - \\ - & r_3 & - \end{pmatrix} \begin{pmatrix} | & | & | \\ a_1 & a_2 & a_3 \\ | & | & | \end{pmatrix} = \begin{pmatrix} r_1a_1 & r_1a_2 & r_1a_3 \\ r_2a_1 & r_2a_2 & r_2a_3 \\ r_3a_1 & r_3a_2 & r_3a_3 \end{pmatrix}.$$

♡ Week 2 Extra ♡

Focus: LU decomposition (aka Gaussian elimination), orthogonal matrices.

LU factorization of a matrix A is a way of writing A as a product of two matrices $\boxed{A = LU}$, where L is a lower-diagonal matrix with units on the diagonal and U is an upper-diagonal matrix.

Definition. A square matrix A is called *orthogonal* if $A^T A = I$. (The unit matrix is denoted by I.)

1. Compute the result of multiplying a row $r = \begin{pmatrix} r_1 & \cdots & r_n \end{pmatrix}$ by a matrix M with rows $M_1, \ldots, M_n$.

Solution:

2. If R is a $k \times n$ matrix and M is a matrix with n rows $M_1, \ldots, M_n$, what are the rows of RM in terms of $M_1, \ldots, M_n$ and the matrix coefficients of R?

Solution:

3. *Bonus. LU factorization = Gaussian elimination.* Solve the system of linear equations using LU factorization:

$$\begin{cases} x + 2y + 3z = 1, \\ y + z = 2, \\ 3x + y - z = 3. \end{cases}$$

Solution:

4. *Not all matrices can be written in LU form.* Show directly why these matrix equations are both impossible (empty spaces mean zeroes):

   a) $\begin{pmatrix} 0 & 1 \\ 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ l & 1 \end{pmatrix} \begin{pmatrix} d & e \\ 0 & f \end{pmatrix};$

   b) $\begin{pmatrix} 1 & 1 & 0 \\ 1 & 1 & 2 \\ 1 & 2 & 1 \end{pmatrix} = \begin{pmatrix} 1 & & \\ l & 1 & \\ m & n & 1 \end{pmatrix} \begin{pmatrix} d & e & g \\ & f & h \\ & & i \end{pmatrix}.$

Solution:

5. *Orthogonal matrices.* Find $A^T A$ if the columns of A are unit vectors, all mutually perpendicular (in this situation, we say that the vectors are *orthonormal*). What if we ask that the rows of A are orthonormal?

Solution:

6. *Bonus.* Say that a square matrix A is factored as a product $A = B^{-1}C$. Perform the same row operation on both B and C, for example add the first row to the second: $B \mapsto B'$, $C \mapsto C'$. Show that this operation does not change the result, that is $A = B^{-1}C = (B')^{-1}C'$.

Hint: think about the first two problems.

Solution:

7. *Binomial formula for matrices.* Show that $(A + B)^2$ is different from $A^2 + 2AB + B^2$ when

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}, \; B = \begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix}.$$

Write down the correct rule: $(A+B)^2 = A^2 + \cdots + B^2$. Can you generalize the rule to $(A+B)^n$?

Solution:

Week 3 Review Session

Focus: rules of matrix multiplication, orthogonal matrices, rotation matrices.

1. When you multiply an $n \times m$ matrix by an $m \times l$ matrix, what are the dimensions of the resulting matrix?

Solution:

2. *Zero scalar vs zero vector vs zero matrix.* Let A be an $n \times m$ matrix, B be an $m \times l$ matrix, v be a column m-vector and r be a row m-vector, for example:

$$A = \begin{pmatrix} 1 & 2 & 0 \\ 1 & 0 & 2 \end{pmatrix}, \; B = \begin{pmatrix} 2 & -2 \\ -1 & 1 \\ -1 & 1 \end{pmatrix}, \; v = \begin{pmatrix} 2 \\ -1 \\ -1 \end{pmatrix}, \; r = \begin{pmatrix} 1 & 2 & 0 \end{pmatrix}.$$

In this case, products AB, Av and rv are all zero, so we can write $AB = 0$, $Av = 0$, $rv = 0$. But would it make sense to write $AB = Av = rv = 0$? Why / why not? Do the results of those operations belong to the same vector space?

Solution:

3. Answer the following questions. Provide explanations.

   a) Is the identity matrix always square? (By the way it can be stored with one parameter.)
   b) Do rectangular matrices have inverses?
   c) Do all square matrices have inverses?
   d) What is the condition for a 2×2 matrix to have inverse?

Solution:

4. *Parallel planes.*

   a) Consider the set of points $\begin{pmatrix} x \\ y \\ z \end{pmatrix}$ in $\mathbb{R}^3$ satisfying condition $2x + 3y + 4z = 0$. Describe this geometric object. Find a normal vector to it.

   b) What if we consider the equation $2x + 3y + 4z = 1$? Why is the normal the same?

Solution:

5. *Row and column operations as matrix multiplication. (Inspired by problem 2.4.8 from Strang.)* Consider the following matrices:

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \; B = \begin{pmatrix} 2 & 0 \\ 0 & -1 \end{pmatrix}, \; C = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \; D = \begin{pmatrix} 0 & 1 \\ 0 & -1 \end{pmatrix}.$$

   How is each row of BA, CA, DA related to the rows of A? How is each column of AB, AC, AD related to the columns of A?

Solution:

6. *(Problem 2.4.5 from Strang.)* Compute A^2 and A^3. Make a prediction for A^5 amd A^n:

   a) $A = \begin{pmatrix} 1 & b \\ 0 & 1 \end{pmatrix}$;

   b) $A = \begin{pmatrix} 2 & 2 \\ 0 & 0 \end{pmatrix}$.

Solution:

7. *Binomial formula for matrices. Matrices do not commute. (Problem 2.4.6 from Strang.)* Show that $(A+B)^2$ is different from $A^2+2AB+B^2$ when

$$A = \begin{pmatrix} 1 & 2 \\ 0 & 0 \end{pmatrix}, \ B = \begin{pmatrix} 1 & 0 \\ 3 & 0 \end{pmatrix}.$$

Write down the correct rule: $(A+B)^2 = A^2 + \cdots + B^2$. Can you generalize the rule to $(A+B)^n$?

Solution:

8. *(Problem 2.4.7 from Strang.)* Is the following true or false? Give counterexamples when false. Matrices A, B and C are such that all the operations are well-defined.

   a) If columns 1 and 3 of B are the same, then so are columns 1 and 3 of AB.

   b) If rows 1 and 3 of B are the same, then so are rows 1 and 3 of AB.

   c) If rows 1 and 3 of A are the same, then so are rows 1 and 3 of ABC.

   d) $(AB)^2 = A^2B^2$.

Solution:

9. *Orthogonal matrices.* Find A^TA if the columns of A are unit vectors, all mutually perpendicular (in this situation, we say that the vectors are *orthonormal*). What if we ask that the rows of A are orthonormal?

Solution:

10. *Defining a matrix by its image. Rotation matrices.* Work out these questions for 2×2 matrices.

   a) If we want a matrix A to send vector $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ to twice itself and vector $\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ to $\left(\begin{smallmatrix}a\\b\end{smallmatrix}\right)$, then what are the matrix entries of A?

   b) If we want a matrix B to send vector $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ $\left(\begin{smallmatrix}a\\b\end{smallmatrix}\right)$ and vector $\left(\begin{smallmatrix}1\\1\end{smallmatrix}\right)$ to twice itself, then what are the matrix entries of B?

   c) What are the coordinates of vectors $\left(\begin{smallmatrix}1\\0\end{smallmatrix}\right)$ and $\left(\begin{smallmatrix}0\\1\end{smallmatrix}\right)$ after we rotate them by the angle θ?

   d) How do you write a matrix that rotates every vector in the plane by the angle θ?

   e) Is the matrix in the previous part orthogonal?

11. *Angles between vectors.* Consider an n-cube C – the set of points in $\mathbb{R}^n$ all of whose coordinates vary from 0 to 1.

   a) In a two-dimensional cube, find the angle between the diagonal and an edge.

   b) In a three-dimensional cube, find the angle between the long diagonal and an edge.

   c) In a three-dimensional cube, find the angle between the long diagonal and a face.

   d) In an n-dimensional cube, find the angle between the long diagonal and an edge.

Recitation 2. February 26

Focus: QR decomposition, SVD, least squares.

1. Say that a square matrix A is factored as a product $A = BC^{-1}$. Perform the same column operation on both B and C, for example add the first column to the second: $B \mapsto B'$, $C \mapsto C'$. Show that this operation does not change the result, that is $A = BC^{-1} = B'(C')^{-1}$.

Solution:

2. *Finding a QR decomposition.* Write the following matrix A as a product $A = QR$ for some orthogonal Q and upper-triangular R:

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}.$$

Solution:

3. *Finding an SVD.* Consider matrix A:

$$A = \begin{pmatrix} 1 & 3 & -1 \\ -2 & -6 & 2 \\ 3 & 9 & -3 \end{pmatrix}.$$

   a) Describe its column space.
   b) Express A as an outer product of two vectors. Is this decomposition unique?
   c) Find a compact and a full form SVD for this matrix.

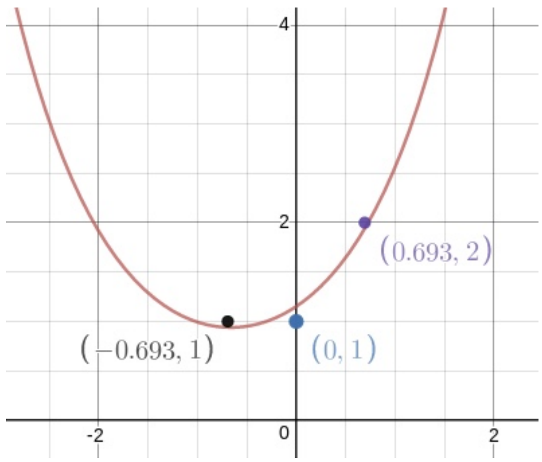
Solution:

Solution:

4. *Least squares approximation.* Consider the set of functions of the form $f(x) = ae^x + be^{-x}$, where a and b vary over real numbers. In this space of functions, use the least squares algorithm to approximate the unknown function U that takes the following values:

$$U\left(\ln\frac{1}{2}\right) = 1,\ U(0) = 1,\ U(\ln 2) = 2.$$

You will get the following picture:



a) Write the sum $S(a, b)$ of squared errors.

b) Write the condition of finding local minimum using partial derivatives with respect to a and b.

c) Write the condition above as a matrix equation and solve this equation.

Solution:

Recitation 2. February 26

Focus: QR decomposition, SVD, least squares.

1. Say that a square matrix A is factored as a product $A = BC^{-1}$. Perform the same column operation on both B and C, for example add the first column to the second: $B \mapsto B'$, $C \mapsto C'$. Show that this operation does not change the result, that is $A = BC^{-1} = B'(C')^{-1}$.

Solution: A column operation on B corresponds to multiplying B on the right by an appropriate invertible matrix, say M. Assuming this rule for a moment, we observe that then $B' = BM$ and $C' = CM$, because we perform the same column operation on both matrices. Therefore, plugging these formulas in, we get $B'(C')^{-1} = (BM)(CM)^{-1}$, and since taking inverses switches the order of factors, we can continue the string of equalities as $B'(C')^{-1} = BMM^{-1}C^{-1} = BC^{-1} = A$. So we achieved the desired result.

In order to understand why multiplying by M performs column operations on B, first write B as a block matrix $B = \begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix}$, where b_i are columns of B. Then we can multiply B by M as block matrices (M is considered to have only trivial 1×1 blocks), and each column of BM would be expressed as a linear combination of columns of B with coefficients from some column of M. For example, if $M = (m_{ij})$, then the first column of BM is:

$$\begin{pmatrix} b_1 & \cdots & b_n \end{pmatrix} \begin{pmatrix} m_{11} \\ \vdots \\ m_{n1} \end{pmatrix} = m_{11}b_1 + m_{21}b_2 + \cdots + m_{n1}b_n.$$

This should have concluded explanation, but it might still look confusing, so let's consider several examples. First, take a 3×3 matrix B written in its block form $B = \begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix}$ and try to find such a matrix M that multiplication by M on the right would add twice the first column to the third. I claim that the following matrix works (blank spaces are zeroes):

$$M = \begin{pmatrix} 1 & 0 & 2 \\ & 1 & 0 \\ & & 1 \end{pmatrix}.$$

Let's check it (don't forget that the b_i are all column 3-vectors):

$$BM = \begin{pmatrix} b_1 & b_2 & b_3 \end{pmatrix} \begin{pmatrix} 1 & 0 & 2 \\ & 1 & 0 \\ & & 1 \end{pmatrix} = \begin{pmatrix} b_1 & b_2 & 2b_1 + b_3 \end{pmatrix}.$$

For the second example, I want to scale up the second column of B by a factor of $\sqrt{5}$. Then I need to take the following matrix:

$$M = \begin{pmatrix} 1 & & \\ & \sqrt{5} & \\ & & 1 \end{pmatrix}.$$

You can check that BM looks like B with its second column scaled by direct block multiplication as above.

Finally, let's say that I want to switch the second and the third column (although this operation will not be used in the remainder of the worksheet). Then I need to take an appropriate *permutation matrix*:

$$M = \begin{pmatrix} 1 & & \\ & 0 & 1 \\ & 1 & 0 \end{pmatrix}.$$

2. *Finding a QR decomposition.* Write the following matrix A as a product $A = QR$ for some orthogonal Q and upper-triangular R:

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}.$$

Solution: In this problem, I want to apply the previous one by writing A as a product AI^{-1}, where I is the identity matrix, and then perform column operations on both A and I simultaneously until the first matrix becomes orthogonal and we get $A = QR$. In order to orthogonalize the first factor, I will use a modified version of Gram-Schmidt process.

Recall that Gram-Schmidt algorithm goes as follows:

- Normalize the first vector.
- Subtract a suitable scalar multiple of the first vector from all the rest vectors so that they become orthogonal to the first vector.
- Normalize the second vector.
- Subtract a suitable scalar multiple of the first vector from the third and later vectors so that they become orthogonal to the second vector. Note that they stay orthogonal to the first vector.
- Etc.

For the purpose of hand calculation, I find it bothersome to normalize vectors before orthogonalizing, because I wouldn't like to carry a lot of square roots around. So I will use the following algorithm for the given 2×2 matrix:

- Subtract a suitable scalar multiple of the first vector from the second so that the result becomes orthogonal to the first vector.
- Normalize the first vector.
- Normalize the second vector.
- Now we have Q explicitly. Compute R.

Note that since we are subtracting multiples of the first column to the second, the second factor becomes upper-triangular as required. More generally, since Gram-Schmidt algorithm subtracts multiples of one column from those to the right, the second factor stays upper-triangular. See the first example in the first problem.

First step. Find a scalar λ such that $\begin{pmatrix} 4 \\ 3 \end{pmatrix} - \lambda \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ is orthogonal to $\begin{pmatrix} 1 \\ 2 \end{pmatrix}$. I claim that $\lambda = \dfrac{\begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 4 \\ 3 \end{pmatrix}}{\begin{pmatrix} 1 \\ 2 \end{pmatrix} \cdot \begin{pmatrix} 1 \\ 2 \end{pmatrix}} = \frac{4+6}{1+4} = 2$ works. Note that dots denote dot product here. Then subtracting the twice the first column from the second in both factor gives us to the following equality (do not forget that the second matrix is inverted!):

$$A = \begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}^{-1}.$$

Now columns of the first factor are orthogonal.

Second step. The magnitude of the first column is $\sqrt{1+2^2} = \sqrt{5}$, so we divide the first column of both matrices by $\sqrt{5}$:

$$A = \begin{pmatrix} 1 & 2 \\ 2 & -1 \end{pmatrix} \begin{pmatrix} 1 & -2 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{\sqrt{5}} & 2 \\ \frac{2}{\sqrt{5}} & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & -2 \\ 0 & 1 \end{pmatrix}^{-1}.$$

Third step. The magnitude of the second column is also $\sqrt{5}$, so we divide the second column of both matrices by $\sqrt{5}$:

$$A = \begin{pmatrix} \frac{1}{\sqrt{5}} & 2 \\ \frac{2}{\sqrt{5}} & -1 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & -2 \\ 0 & 1 \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ 0 & \frac{1}{\sqrt{5}} \end{pmatrix}^{-1}.$$

Fourth step. Compute the inverse. For that, we first compute the quatity $ad - bc$, because we will need to divide by it. So $ad - bc = \frac{1}{\sqrt{5}} \cdot \frac{1}{\sqrt{5}} - 0 = \frac{1}{5}$. Dividing by it means multiplying by 5. So apply our formula for inverses of 2×2 matrices:

$$A = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{5}} \\ 0 & \frac{1}{\sqrt{5}} \end{pmatrix}^{-1} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \end{pmatrix} 5 \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ 0 & \frac{1}{\sqrt{5}} \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{-1}{\sqrt{5}} \end{pmatrix} \begin{pmatrix} \sqrt{5} & 2\sqrt{5} \\ 0 & \sqrt{5} \end{pmatrix}.$$

3. *Finding an SVD.* Consider matrix A:

$$A = \begin{pmatrix} 1 & 3 & -1 \\ -2 & -6 & 2 \\ 3 & 9 & -3 \end{pmatrix}.$$

a) Describe its column space.

b) Express A as an outer product of two vectors. Is this decomposition unique?

c) Find a compact and a full form SVD for this matrix.

Solution:

a) $\text{Col}(A) = \text{Span}\begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix}$, because all other columns are scalar multiples of the first one.

b) $A = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} \begin{pmatrix} 1 & 3 & -1 \end{pmatrix}$. The decomposition is not unique, because for example, we can multiply the column by 5 and divide the row by 5, and the result will not change.

c) A compact form SVD follows almost immediately from the column-row decomposition – what remains is to normalize the row and teh column. Note that (1) denotes a 1×1 matrix as opposed to a scalar 1.

$$A = \begin{pmatrix} 1 \\ -2 \\ 3 \end{pmatrix} \begin{pmatrix} 1 \end{pmatrix} \begin{pmatrix} 1 & 3 & -1 \end{pmatrix} = \begin{pmatrix} \frac{1}{\sqrt{14}} \\ \frac{-2}{\sqrt{14}} \\ \frac{3}{\sqrt{14}} \end{pmatrix} \begin{pmatrix} \sqrt{154} \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{11}} & \frac{3}{\sqrt{11}} & \frac{-1}{\sqrt{11}} \end{pmatrix}.$$

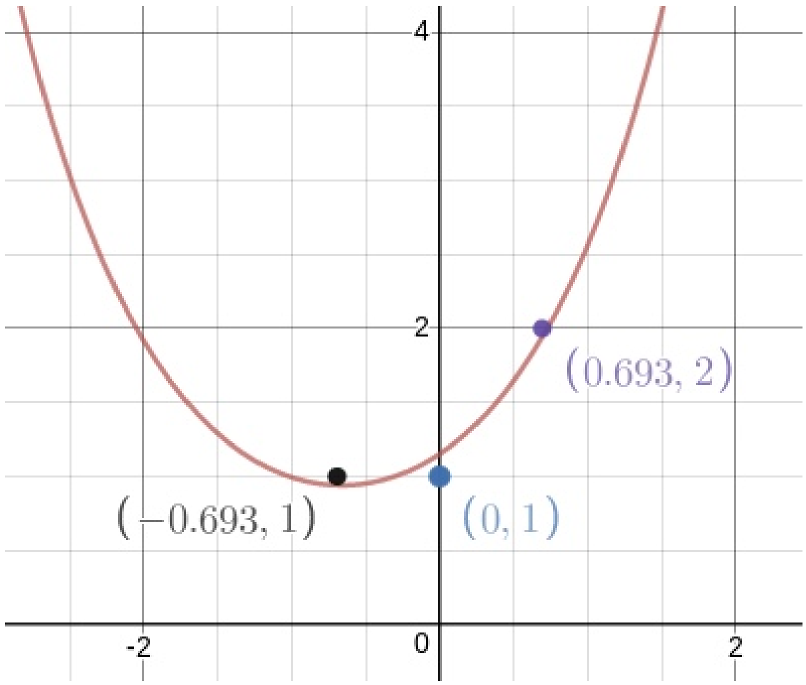
A full form SVD can be found by picking two complementary orthogonal vector to the column, and then to the row. Here is an example for the column:

$$A = \begin{pmatrix} \frac{1}{\sqrt{14}} & \frac{1}{\sqrt{3}} & \frac{5}{\sqrt{42}} \\ \frac{-2}{\sqrt{14}} & \frac{-1}{\sqrt{3}} & \frac{4}{\sqrt{42}} \\ \frac{3}{\sqrt{14}} & \frac{-1}{\sqrt{3}} & \frac{1}{\sqrt{42}} \end{pmatrix} \begin{pmatrix} \sqrt{154} & & \\ & 0 & \\ & & 0 \end{pmatrix} \begin{pmatrix} \frac{1}{\sqrt{11}} & \frac{3}{\sqrt{11}} & \frac{-1}{\sqrt{11}} \\ * & * & * \\ * & * & * \end{pmatrix}.$$

4. *Least squares approximation.* Consider the set of functions of the form $f(x) = ae^x + be^{-x}$, where a and b vary over real numbers. In this space of functions, use the least squares algorithm to approximate the unknown function U that takes the following values:

$$U\left(\ln\frac{1}{2}\right) = 1,\; U(0) = 1,\; U(\ln 2) = 2.$$

You will get the following picture:



a) Write the sum $S(a, b)$ of squared errors.

b) Write the condition of finding local minimum using partial derivatives with respect to a and b.

c) Write the condition above as a matrix equation and solve this equation.

Solution:

a) $S(a, b) = \sum_{i=1}^{n} (f(x_i) - y_i)^2$. In our case, $n = 3$ and the datapoints are given by the values of U, so we plug this in: $S(a, b) = \left(\frac{1}{2}a + 2b - 1\right)^2 + (a + b - 1)^2 \left(2a + \frac{1}{2}b - 2\right)^2$.

b) First compute partial derivative with respect to a:

$$\frac{\partial S}{\partial a}(a, b) = 2\sum_{i=1}^{n} \left(ae^{x_i} + be^{-x_i} - y_i\right) e^{x_i} = 2\sum_{i=1}^{n} \left(ae^{2x_i} + b - y_i e^{x_i}\right).$$

Combine similar summands:

$$\frac{\partial S}{\partial a}(a, b) = 2\left(\sum_{i=1}^{n} e^{2x_i}\right) a + 2\left(\sum_{i=1}^{n} 1\right) b - 2\left(\sum_{i=1}^{n} y_i e^{x_i}\right).$$

Now we can plug in our datapoints:

$$\frac{\partial S}{\partial a}(a, b) = \frac{21}{4}a + 3b - \frac{11}{2}.$$

Similarly for b:

$$\frac{\partial S}{\partial b}(a, b) = 2\left(\sum_{i=1}^{n} 1\right) a + 2\left(\sum_{i=1}^{n} e^{-2x_i}\right) b - 2\left(\sum_{i=1}^{n} y_i e^{-x_i}\right).$$

And now with the given data:

$$\frac{\partial S}{\partial b}(a,b) = 3a + \frac{21}{4}b - 4.$$

The condition for finding an extremum is for all the partial derivatives to vanish, so:

$$\begin{cases} \frac{21}{4}a + \ \ 3b = \frac{11}{2}, \\ \ 3a + \frac{21}{4}b = 4. \end{cases}$$

c) $\begin{pmatrix} \frac{21}{4} & 3 \\ 3 & \frac{21}{4} \end{pmatrix} \begin{pmatrix} a \\ b \end{pmatrix} = \begin{pmatrix} \frac{11}{2} \\ 4 \end{pmatrix}$. Solution can be obtained by multiplying both sides by the inverse of $\begin{pmatrix} \frac{21}{4} & 3 \\ 3 & \frac{21}{4} \end{pmatrix}$ on the left. The resulting coefficients are $a = \frac{30}{33}$ and $b = \frac{8}{33}$. You can find the plot of the corresponding function above.

Recitation 3. March 5

Focus: exam comments, linear independence, projections.

Definition. Recall that if A is a matrix with independent columns, then the *projection matrix* P on the column space of A can be written as $P = A(A^TA)^{-1}A^T$.

1. *Recognizing vector spaces.* Is the following set a vector space?

   a) The set of all vectors in $\mathbb{R}^3$ except those of the form $\begin{pmatrix} x & 0 & 0 \end{pmatrix}^T$ with $x > 0$.
   b) The set of 2×3 matrices whose six elements sum to 6.
   c) The set of rank one 3×3 matrices together with the zero matrix.

   Solution:

2. *Vector spaces and bases.* Let V be the space of homogeneous quadratic polynomials in two variables, i.e. polynomials of the form $f(x, y) = ax^2 + bxy + cy^2$.

   a) Are elements x^2, xy, y^2 linearly independent?
   b) What about x^2, $x^2 + xy + y^2$, $xy + y^2$?
   c) And x^2, $x^2 + xy$, $x^2 + xy + y^2$?
   d) What is the dimension of this vector space V?

   Solution:

3. *Subspaces.* Let V be the space from the previous problem. Consider the subset of V that consists of functions $f \in V$ such that $f(1, 1) = 0$. Denote this subset by W – it is a vector subspace of V.

   a) Prove that W is a vector space.
   b) Find a basis of this vector space.
   c) What is the dimension of W?

   Solution:

4. *Projection onto a subspace.* Consider the following matrix A written as a full SVD $A = U\Sigma V^T$:

$$A = \begin{pmatrix} \frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & 0 & -\frac{1}{\sqrt{2}} \\ 0 & 1 & 0 \end{pmatrix} \begin{pmatrix} 10 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 \end{pmatrix}^T.$$

a) What is the rank of A?

b) Mark U_1, U_2, V_1 and V_2.

c) Circle columns of U that span $\operatorname{col} A$.

d) Compute the projection matrix on $\operatorname{col} A$.

Solution:

5. *Bonus. Might be useful for PSet 4. (Or not.)* If A is decomposed into a product $A = BC$ with C being square invertible, then $\operatorname{col} A = \operatorname{col} B$.

6. *Bonus.* We declare x^2, xy and y^2 to be the standard orthonormal basis in V, that is we write them as follows:

$$x^2 = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix}, \ xy = \begin{pmatrix} 0 \\ 1 \\ 0 \end{pmatrix}, \ y^2 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix}.$$

Compute the projection matrix on W.