Homework Week #2

Ho Chi Vuong
AI Math Foundations: Abstract Vector Spaces
Center of Talent in AI

April 12, 2020

Section P

P2.1

The last (seems redundant) line allows us to combine any number of affine transformations (e.g., a series of rotation, translation, then rotation) into one by multiplying the respective matrices. Verify this for yourself!

tl;dr: The last line in the augmented matrix $0, \ldots, 1$ is used to aid in the multiplication of matrices. Without it, we have undefined results in our matrix multiplication

Let's say we have a series of two affine transformation, wrapped in two augmented matrice

$$T_{t_1} = \begin{bmatrix} 1, 0, 0, p_1 \\ 0, 1, 0, p_2 \\ 0, 0, 1, p_2 \\ 0, 0, 0, 1 \end{bmatrix}, T_{t_2} = \begin{bmatrix} 1, 0, 0, p_3 \\ 0, 1, 0, p_4 \\ 0, 0, 1, p_5 \\ 0, 0, 0, 1 \end{bmatrix}$$

When combining in series of affine transformations, we have the following matrix multiplication:

$$T_{t_1} T_{t_2} = \begin{bmatrix} 1 & 0 & 0 & p_1 \\ 0 & 1 & 0 & p_2 \\ 0 & 0 & 1 & p_3 \\ 0 & 0 & 0 & 1 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & p_3 \\ 0 & 1 & 0 & p_4 \\ 0 & 0 & 1 & p_5 \\ 0 & 0 & 0 & 1 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & p_1 + p_3 \\ 0 & 1 & 0 & p_2 + p_4 \\ 0 & 0 & 1 & p_3 + p_5 \\ 0 & 0 & 0 & 1 \end{bmatrix}$$

allowing us combine any number of affine transformation. If the line $0, \dots, 1$ is not there, we have the following matrix multiplication:

$$T_{t_1}T_{t_2} = \begin{bmatrix} 1 & 0 & 0 & p_1 \\ 0 & 1 & 0 & p_2 \\ 0 & 0 & 1 & p_3 \end{bmatrix} \cdot \begin{bmatrix} 1 & 0 & 0 & p_3 \\ 0 & 1 & 0 & p_4 \\ 0 & 0 & 1 & p_5 \end{bmatrix} = undefined$$

P2.2

P2.3

Prove that Matrix multiplication $T : \mathbb{R}^{n \times p} \rightarrow \mathbb{R}^{m \times p}$ with $TC = D, T = B_{m \times n}$ is a linear map. Find matrix representation of this linear transformation in the standard bases.

To prove $T : \mathbb{R}^{n \times p} \rightarrow \mathbb{R}^{m \times p}$ with $TC = D$, $T = B_{m \times n}$ is a linear transformation, we have to prove:

1. For $x, x' \in \mathbb{R}^{n \times p}$, $T(x + x') = Tx + Tx'$
2. For $x \in \mathbb{R}^{n \times p}$ and $c \in \mathbb{R}$, $T(cx) = cTx$

1. We have

$$T(x + x') = B_{m \times n}(x + x') = B_{m \times n} \cdot \begin{bmatrix} x_{11} + x'_{11} & \dots & x_{n1} + x'_{n1} \\ \dots & \ddots & \dots \\ x_{1p} + x'_{1p} & \dots & x_{np} + x'_{np} \end{bmatrix}$$

and

$$Tx + Tx' = B_{m \times n} \cdot \begin{bmatrix} x_{11} & \dots & x_{n1} \\ \dots & \ddots & \dots \\ x_{1p} & \dots & x_{np} \end{bmatrix} + B_{m \times n} \cdot \begin{bmatrix} x'_{11} & \dots & x'_{n1} \\ \dots & \ddots & \dots \\ x'_{1p} & \dots & x'_{np} \end{bmatrix}$$

making $T(x + x') = Tx + Tx'$ true.

2. We have

$$T(cx) = B_{m \times n} \cdot \begin{bmatrix} cx_{11} & \dots & cx_{n1} \\ \dots & \ddots & \dots \\ cx_{1p} & \dots & cx_{np} \end{bmatrix}$$

and

$$cTx = cB_{m\times n} \cdot \begin{bmatrix} x_{11} & \dots & x_{n1} \\ \dots & \ddots & \dots \\ x_{1p} & \dots & x_{np} \end{bmatrix} = B_{m\times n} \cdot \begin{bmatrix} cx_{11} & \dots & cx_{n1} \\ \dots & \ddots & \dots \\ cx_{1p} & \dots & cx_{np} \end{bmatrix}$$

making $c \in R$, $T(cx) = cTx$ true.
With 1 and 2 we can prove the above.

P2.4

Prove that for linear map f,the $\ker(f)$ is a subspace of $\mathbb{R}^5$ with $f : \mathbb{R}^5 \mapsto \mathbb{R}^5$. If $f(x_1, ..., x_5) = (x_1, ..., x_4, 0)$ then is f a linear function? Find the matrix of f, $\ker(f)$ and $\dim(\ker)$. Note that f is a projection.

f is of course a linear function because it is itself a map between two vector spaces that preserves addition and scalar multiplication. We have the following:

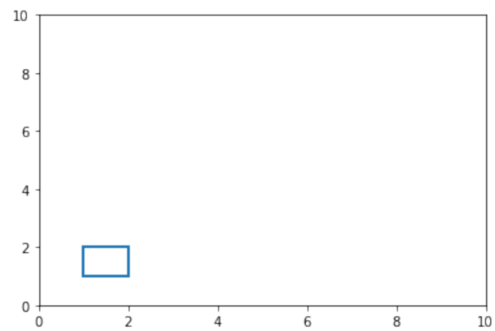
1. For $x, x' \in \mathbb{R}^5$, $T(x+x') = T((x_1+x'_1, \dots, x_5+x'_5) = (x_1+x'_1, \dots, 0) = (x_1, \dots, 0) + (x'_1, \dots, 0) = Tx + Tx'$
2. For $x \in \mathbb{R}^5$ and $c \in \mathbb{R}$, $T(cx) = T((cx_1, \dots, cx_5)) = (cx_1, \dots, 0) = c(x_1, \dots, 0) = cTx$

The matrix of f is $\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$.
$ker(f) = \{(0, \dots, x_5) \in \mathbb{R}^5 : x_5 \in \mathbb{R}\}$, and as the basis of $ker(f)$ is $(0, 0, 0, 0, 1)$, $dim(ker) = 1$.

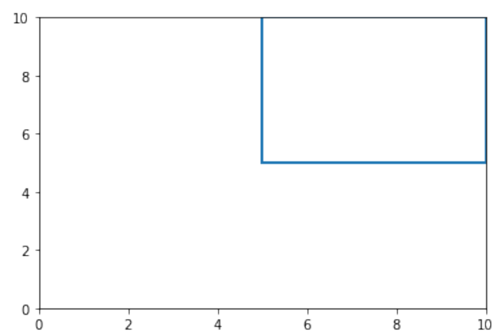
P2.5

Verify for yourself the effect of coordinate scaling Sx with scaling or diagonal matrix$S = \mathsf{diag}(\lambda_1, \dots, \lambda_n) \in \mathbb{R}^{n\times n}$.

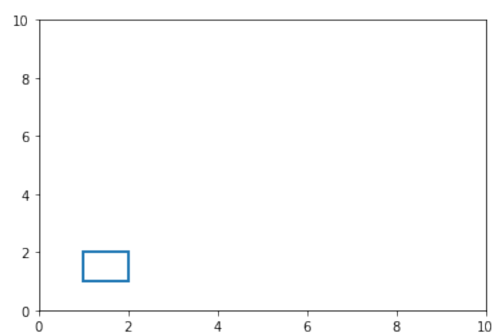
I have visualized it down here for the sake of comprehension.
Sx with $S = \mathsf{diag}(5, 5)$
Before:



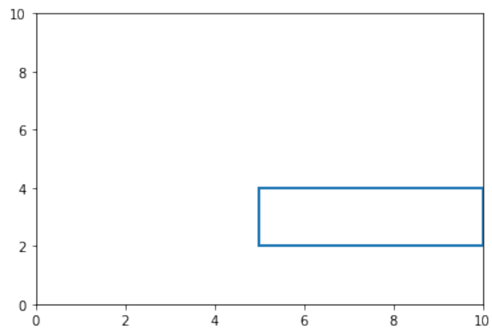
After:



Sx with $S = \mathsf{diag}(5, 2)$



After:



P2.8

In PCA: input image $X \overset{T}{\rightleftharpoons}$ Coordinate vector $\bar{z} = (1, w_1, \ldots, w_n)^\top$ (no reduction). Then with dimensionality reduction: $\bar{z} \xrightarrow{A} z = (1, w_1, \ldots, w_k)^\top$, i.e., keeping only first $k+1$ coordinates. Find matrix A of the coordinate transformation.

$$A = \begin{bmatrix} 1 & 0 & \ldots & 0 & \ldots & 0 \\ 0 & 1 & \ldots & 0 & \ldots & 0 \\ \vdots & \vdots & \ddots & 0 & \ldots & 0 \\ 0 & 0 & 0 & 1 & \ldots & 0 \end{bmatrix}$$

or in other words, A is a $(k+1) \times n$ matrix with $a_{ii} = 1$, $i = 1, 2, \ldots, k+1$, and the rest equals 0

P2.9

Prove that for any A: $\mathsf{rank}(A^\top A) = \mathsf{rank}(AA^\top) = \mathsf{rank}(A)$.
Let's first prove $\mathsf{rank}(AA^\top) = \mathsf{rank}(A)$.
For $x \in N(A)$, we have:

$$Ax = 0$$

$$\Rightarrow A^\top Ax = 0$$

$$\Rightarrow x \in N(A^\top A)$$

therefore,

$$N(A) \subseteq N(A^\top A)$$

Again, for $x \in N(A^\top A)$, we have:

$$A^\top Ax = 0$$
$$\Rightarrow x^\top A^\top Ax = 0$$
$$\Rightarrow (Ax)^\top (Ax) = 0$$
$$\Rightarrow Ax = 0$$

therefore,

$$N(A^\top A) \subseteq N(A)$$

Thus,

$$N(A^\top A) = N(A)$$
$$\Rightarrow \mathsf{nullity}(A^\top A) = \mathsf{nullity}(A)$$
$$\Rightarrow \mathsf{rank}(A^\top A) = \mathsf{rank}(A)$$

Then we prove $\mathsf{rank}(A^\top A) = \mathsf{rank}(AA^\top)$.
If $n < m$, we have:

$$\mathsf{rank}(AA^\top) = \mathsf{rank}\left(\begin{bmatrix} AA^\top & \mathbf{0}_{n\times(m-n)} \\ \mathbf{0}_{(m-n)\times n} & \mathbf{0}_{(m-n)\times(m-n)} \end{bmatrix}\right)$$
$$= \mathsf{rank}\left(\begin{bmatrix} A \\ \mathbf{0}_{(m-n)\times m} \end{bmatrix} \begin{bmatrix} A \\ \mathbf{0}_{(m-n)\times m} \end{bmatrix}^\top\right)$$
$$= \mathsf{rank}\left(\begin{bmatrix} A \\ \mathbf{0}_{(m-n)\times m} \end{bmatrix}\right)$$
$$= \mathsf{rank}(A)$$

Otherwise, if $m < n$, we have:

$$\mathsf{rank}(AA^t) = \mathrm{rank}\left(\begin{bmatrix} A & \mathbf{0}_{n\times(n-m)} \end{bmatrix} \begin{bmatrix} A & \mathbf{0}_{n\times(n-m)} \end{bmatrix}^\top\right)$$
$$= \mathsf{rank}\left(\begin{bmatrix} A & \mathbf{0}_{n\times(n-m)} \end{bmatrix}\right)$$
$$= \mathsf{rank}(A)$$

P2.11

Show that for square & full rank matrices:

$$(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$$
$$(A^{-1})^T = (A^T)^{-1} =: A^{-\top}$$

To prove $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$, we have the following:

$$(ABC)C^{-1}B^{-1}A^{-1} = AB(CC^{-1})B^{-1}A^{-1} = ABB^{-1}A^{-1} = AA^{-1} = I$$

From this we can conlude that the inverse of ABC is $C^{-1}B^{-1}A^{-1}$, which also means $(ABC)^{-1} = C^{-1}B^{-1}A^{-1}$.
To prove $(A^{-1})^T = (A^T)^{-1} =: A^{-\top}$, we have the following:

$$A^\top(A^{-1})^\top = (A^{-1}A)^\top = I = A^\top(A^\top)^{-1}$$
$$\Rightarrow (A^{-1})^\top = (A^\top)^{-1}$$

P2.11

Prove again for yourself $\mathsf{rank}(A_{m\times n}) = \mathsf{rank}(A^\top) \leq \min(m,n)$
Since $\mathsf{rank}(A_{m\times n})$ = the number of linearly independent rows/columns, which is also $\leq min(m,n)$, transposing the matrix which results in $A_{n\times m}$ will change nothing to this definition. That is, $\mathsf{rank}(A^\top) = \mathsf{rank}(A_{m\times n})$ = the number of linearly independent rows/columns, which is also $\leq min(m,n)$, is also equals to $\mathsf{rank}(A_{m\times n})$, since the linear dependence property does not change when transposing matrix.

Section E

E2.1

Let $\beta = \{v_1, \ldots, v_n\}$ be a basis of a vector space V, and $T : V \to W$ be a linear transformation from V to W. Is the following proposition true? Why? If the formula is wrong, correct it.

$$T(\sum_{i=1}^{n} a_i v_i) = T(v) \Leftrightarrow \sum_{i=1}^{n} a_i v_i = v$$

The proposition is correct. With T as a linear function, we have the following:

$$T(\sum_{i=1}^{n} a_i v_i) = \sum_{i=1}^{n} a_i T v_i = v = \sum_{i=1}^{n} a_i w_i = w = T(v)$$

with

$$\sum_{i=1}^{n} a_i v_i = v$$

E2.2

Let V be a finite-dimensional vector space and $\alpha = \{v_1, \ldots, v_n\}$ be a basis of V. Let W be another vector space with some vectors $\beta = \{w_1, \ldots, w_n\}$. Prove that there exists exactly one linear transformation $T : V \to W$ such that $T(v_j) = w_j, \forall j$.
We have to prove two things, that

1. there exists such linear transformation T, and
2. there is at most one such transformation.

1. Assume for contradiction that we have $T : V \to W, U : V \to W$, $T(v_j) = w_j$ and $U(v_j) = w_j$ for $j = 1, \ldots, n$. We have any $v \in V$ with the following:

$$v = a_1 v_1 + \ldots + a_n v_n$$

Since T is linear, we have:

$$Tv = a_1 T v_1 + \ldots + a_n T v_n$$

equals to

$$Tv = a_1 w_1 + \ldots + a_n w_n$$

The same applies to Uv, as in the following:

$$Uv = a_1 w_1 + \ldots + a_n w_n$$

As above, we have:

$$Tv = Uv \Leftrightarrow T = U$$

contradicting the original assumptions, proving (1)
2.We have the above in (1):

$$Tv = a_1w_1 + \ldots + a_nw_n$$

and

$$v_j = 0v_1 + \ldots + 1v_j + \ldots + 0v_n$$

giving us

$$Tv_j = 0w_1 + \ldots + 1w_j + \ldots + 0w_n = w_j$$

We also have to prove that T is linear, which is true because for any $v, v' \in V$ and $c \in \mathbb{R}$, we have the following:

$$cTv + cTv' = c(a_1w_1 + \ldots + a_nw_n) + c(b_1w_1 + \ldots + b_nw_n)$$

$$= c(a_1 + b_1)w_1 + \ldots + c(a_n + b_n)w_n = T(cv + cv')$$

Thus we can prove (2). With (1),(2), we can prove the above.

E2.3

What I intuitively think of matrix multiplication is that the result of the calculation is the matrix containing the linear combination of, say AB, a row of A and column of B. We can also think of matrix multiplication as the composition of *linear functions*, as matrix can represent T in, for example, $T : V \to W$.

E2.4

Prove that the null space and the row space of a matrix are orthogonal, i.e. every vector in null space is orthogonal to every vector in row space (zero dot product).

Assume we have a matrix $A = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$, its null space $N(A)$ which contains vectors x satisfying $Ax = 0$, and its row space $R = span(v_1, \ldots, v_n)$. We have $Ax = 0 \Leftrightarrow v_1x = 0, \ldots, v_nx = 0$. Therfore, for $r \in R$, we have the following:

$$rx = span(v_1, \ldots, v_n)x = a_1v_1x + \ldots + a_nv_nx = 0$$

From here we have proven that $rx = 0 \forall r \in R$, which means every r is ortogonal with x, which proves what we have to prove.

Assume we have a matrix $A = \begin{bmatrix} v_1 \\ \vdots \\ v_n \end{bmatrix}$, its null space $N(A)$ which contains vectors x satisfying $Ax = 0$, and its row space $R = span(v_1, \ldots, v_n)$. We have $Ax = 0 \Leftrightarrow v_1 x = 0, \ldots, v_n x = 0$. Therfore, for $r \in R$, we have the following:

$$rx = span(v_1, \ldots, v_n)x = a_1 v_1 x + \ldots + a_n v_n x = 0$$

From here we have proven that $rx = 0 \forall r \in R$, which means every r is ortogonal with x, which proves what we have to prove.

E2.5

How can we find the null space and the column space of a matrix ? Write a pseudocode for your algorithm.

Pseudocode for finding null space (incomplete):
function ReducedEchelonForm(A):
..........for i in size of A:
....................multiply all elements of all A[k] (with k not equal i) with GCD(A[i][i], A[k][i])
....................minus all elements of all A[k] (with k not equal i) with (all elements of
A[i]*GCD(A[i][i], A[k][i]))
....................divide all elements of A[i] with A[i][i]