

TP-2 : Linear algebra with matrices and vectors

Note : I'm aware the PDF version of the rapport is cumbersome to read because of the image cuts. I couldn't figure out a way to fix this issue. I vastly recommend to open the rapport via the .md file into VS code and use a previewer to see its content over using the PDF, even if it's more stable.

Ex 1

Part 1 : Practice with Numbers

If my writing is too bad to read, please reach out to me, and I'll happily clarify everything !

Problem 1 to 7 :

Ex TP2

Part 1 : Matrices

Ex 1.1

1)
$$\begin{bmatrix} 5 & 2 \\ 3 & 4 \\ 1 & 6 \end{bmatrix} + \begin{bmatrix} 10 & 50 \\ 20 & 30 \\ 30 & 20 \end{bmatrix} = \begin{bmatrix} 15 & 52 \\ 23 & 34 \\ 31 & 26 \end{bmatrix}$$

2)
$$\begin{bmatrix} 5 & 4 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = 5 \cdot 3 + 4 \cdot 4 = 15 + 16 = 31$$

3)
$$\begin{bmatrix} 3 & 5 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 3 \cdot 3 + 5 \cdot 4 \\ 3 \cdot 0 + 5 \cdot 1 \end{bmatrix} = \begin{bmatrix} 29 \\ 5 \end{bmatrix}$$
 this is a 1×2 matrix \times a 2×2 matrix = 1×2 out

4)
$$\begin{bmatrix} 2 & 5 \\ 40 & 20 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 4 & 1 \end{bmatrix} = \begin{bmatrix} 2 \cdot 3 + 5 \cdot 4, 0 \cdot 3 + 1 \cdot 5 \\ 3 \cdot 40 + 4 \cdot 20, 0 \cdot 4 + 1 \cdot 20 \end{bmatrix} = \begin{bmatrix} 26 & 5 \\ 200 & 20 \end{bmatrix}$$
 $2 \times 2 \cdot 2 \times 2$

5)
$$\begin{bmatrix} 5 & 3 & 7 \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} \Rightarrow \text{not possible}$$
 $1 \times 3 \cdot 2 \times 1$

$$6) \begin{bmatrix} 3 \\ 5 \\ 4 \end{bmatrix} \begin{bmatrix} 5 & 3 & 7 \end{bmatrix} = \begin{bmatrix} 5 \cdot 3 & 3 \cdot 3 & 7 \cdot 3 \\ 5 \cdot 4 & 3 \cdot 4 & 7 \cdot 4 \end{bmatrix} = \begin{bmatrix} 15 & 9 & 21 \\ 20 & 12 & 28 \end{bmatrix}$$

$2 \times 1 \times 3$

$$7) \begin{bmatrix} 0 & 1 & 2 \\ 10 & 10 & 5 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 0 \cdot 1 + 4 \cdot 1 + 7 \cdot 2, 1 \cdot 0 + 5 \cdot 1 + 8 \cdot 2, 3 \cdot 0 + 6 \cdot 1 + 9 \cdot 2 \\ 1 \cdot 10 - 4 \cdot 10 + 7 \cdot 5, 2 \cdot 10 - 5 \cdot 10 + 8 \cdot 5, 3 \cdot 10 - 6 \cdot 10 + 9 \cdot 5 \end{bmatrix}$$

$2 \times 3 \sqrt{3 \times 3}$

$$= \begin{bmatrix} 18 & 21 & 24 \\ 5 & 10 & 15 \end{bmatrix}$$

Problem 8 to 14 :

TP 2 - linear alg

$\mathbb{E} \times \mathbb{I}$ suite

Pari 1 suite

$$8) \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} 7 & 8 & 9 \\ 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix}$$

$3 \times 3 \sqrt{3 \times 3}$

$$9) 5 \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} = \begin{bmatrix} 5 & 10 \\ 15 & 20 \\ 25 & 30 \end{bmatrix}$$

$$10) -1 \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ -5 & -6 \end{bmatrix} = \begin{bmatrix} -1 & -2 \\ -3 & -4 \\ 5 & 6 \end{bmatrix}$$

$$11) \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 5 & 6 \end{bmatrix} - \begin{bmatrix} 10 & 20 \\ 30 & 40 \\ 50 & 60 \end{bmatrix} = \begin{bmatrix} 1-10 & 2-20 \\ 3-30 & 4-40 \\ 5-50 & 6-60 \end{bmatrix} = \begin{bmatrix} -9 & -18 \\ -27 & -36 \\ -45 & -54 \end{bmatrix}$$

$$12) \begin{bmatrix} 2 & 1 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}' = ?$$

$$13) \begin{bmatrix} 1 & 6 \\ 3 & 4 \\ 5 & 2 \end{bmatrix}^T = \begin{bmatrix} 1 & 3 & 5 \\ 6 & 4 & 2 \end{bmatrix}$$

$$14) \begin{pmatrix} 5 & 2 \\ 3 & 9 \\ 1 & 6 \end{pmatrix}^T = \begin{bmatrix} 5 & 3 & 1 \\ 2 & 4 & 6 \end{bmatrix}^T = \begin{bmatrix} 5 & 2 \\ 3 & 4 \\ 1 & 6 \end{bmatrix}$$

For exercise 12, I left it blank because I wasn't sure what the ' prime symbol meant at the time, since we hadn't seen that notation in our lessons. I know the rest of the group generally assumed it was a transpose, and I share this interpretation, so I'll write the transposed matrix just in case :

$$\begin{bmatrix} 2 & 1 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}' = \begin{bmatrix} 2 & 1 \\ 3 & 4 \\ 5 & 6 \end{bmatrix}^T = \begin{bmatrix} 2 & 1 & 5 \\ 1 & 4 & 6 \end{bmatrix}$$

Part 2 : Matrix type

1. Square matrix

- i. A square matrix has the same number of columns as it has of rows.

2. Column matrix

- i. A column matrix has dimension $n \times 1$, so it only has one column. It is vertical => Tall

3. Row matrix

- i. A row matrix has dimension $1 \times m$, so it has many columns but only one row => Wide

4. Identity matrix

- i. The identity matrix has its diagonal element all at 1 (when $n = m$), and the rest of its element at 0 (when $n \neq m$). It is **always** square.

- ii. In a way, it is a sort of permutation matrix. It permutes the vector into itself ?

- iii. If we multiply it with a vector, we get the vector itself.

- iv. The identity matrix is a stricter diagonal matrix.

5. Diagonal matrix

- i. The diagonal matrix has a similar structure to the identity matrix, in a way : when $n = m$ (its diagonal), the element is any number but 0. When not, the number is 0.

6. symmetric matrix

- i. A symmetric matrix has a symmetry access along its diagonal (when $n = m$). So, the

element at n, m is the same than the element at m, n

7. Permutation matrix

- i. A permutation matrix is a matrix that has only one 1 per row **AND** per column.
- ii. The rest of the elements are 0.
- iii. It is always square

8. Upper triangular matrix

- i. An upper triangular matrix is a matrix where the elements below the diagonal (when $n > m$) are all 0.
- ii. The diagonal itself can be any number.
- iii. The above part can be any number as well.

9. Lower triangular matrix

- i. A lower triangular matrix is a matrix where the elements above (when $n < m$) the diagonal are all 0.
- ii. The elements in the diagonal and below can be any number

Part 3 : Some matrix Operations

For

$$1. \frac{1}{n} \mathbf{1} \mathbf{1}^T$$

This result in a matrix with dimension $n \times n$, assuming that the 1-vectors have n instances, where all elements are $\frac{1}{n}$

$$2. \frac{1}{n} \mathbf{1} \mathbf{1}^T \mathbf{x}$$

Here, we're doing a multiplication between the $n \times n$ matrix $\frac{1}{n} \mathbf{1} \mathbf{1}^T$ and one $n \times 1$ vector \mathbf{x} . This result in a $n \times 1$ vector where every element x_i is the mean of \mathbf{x} , μ_x .

$$3. \mathbf{x} - \frac{1}{n} \mathbf{1} \mathbf{1}^T \mathbf{x}$$

Here, we'll be taking our constant mean of \mathbf{x} vector we computed earlier, and subtracting it from the base \mathbf{x} . In essence, it means that for each element \mathbf{x}_i , we subtract μ_x . This result in a vector $n \times 1$ based on \mathbf{x} that has been demeaned, so it has a new mean of 0.

$$4. [\mathbf{I} - \frac{1}{n} \mathbf{1} \mathbf{1}^T] \mathbf{x}$$

This notation is equivalent to the precedent notation, but arrive to the result in a different, arguably more elegant way. First, $[\mathbf{I} - \frac{1}{n} \mathbf{1} \mathbf{1}^T]$ is called a centering matrix. It is a $n \times n$ matrix where every element on the diagonal is $1 - \frac{1}{n}$, and every other element is $\frac{1}{n}$. As its name suggest, it center the vector around its mean.

We can think of it this way : the identity matrix \mathbf{I} preserve the vector as it is. $\frac{1}{n} \mathbf{1} \mathbf{1}^T$ is a matrix that, when multiplying a vector, turn it into its mean. So, when we distribute the vector \mathbf{x} , we arrive to this situation :

$$\mathbf{I}\mathbf{x} - \frac{1}{n} \mathbf{1} \mathbf{1}^T \mathbf{x}$$

which can be simplified to this since $\mathbf{I}\mathbf{x} = \mathbf{x}$:

$$\mathbf{x} - \frac{1}{n} \mathbf{1} \mathbf{1}^T \mathbf{x}$$

We've seen that $\frac{1}{n} \mathbf{1} \mathbf{1}^T \mathbf{x}$ is the mean vector, so we can more clearly see what we're subtracting. And as vector subtraction is done element wise, we're doing this $(x_i - \mu_x)$ to each elements of \mathbf{x} .

5. $[\mathbf{I} - \frac{1}{n} \mathbf{1} \mathbf{1}^T] \mathbf{1}$

Instead of taking the vector \mathbf{x} here, we use the 1-vector. Because this vector is already constituted by its means (all elements are 1, the mean is 1), the resulting vector \mathbf{x}' is the $\mathbf{0}$ -vector. To see more clearly what we're doing :

$$[\mathbf{I} - \frac{1}{n} \mathbf{1} \mathbf{1}^T] \mathbf{1} = \mathbf{I} \mathbf{1} - \frac{1}{n} \mathbf{1} \mathbf{1}^T \mathbf{1} = \mathbf{1} - \frac{1}{n} \mathbf{1} \mathbf{1}^T \mathbf{1}$$

Since $\frac{1}{n} \mathbf{1} \mathbf{1}^T \mathbf{1}$ is the operation to obtain the mean of the vector in red, and that the mean of every elements of the 1-vector (in red) is 1, we arrive in this position :

$$\frac{1}{n} \mathbf{1} \mathbf{1}^T \mathbf{1} = \mathbf{1}$$

Thus leading to have this situation :

$$\mathbf{1} - \frac{1}{n} \mathbf{1} \mathbf{1}^T \mathbf{1} = \mathbf{1} - \mathbf{1} = \mathbf{0}$$

This is the same thing to what we applied in the 4. exercice, except this time the vector is a constant vector (a constant vector is one where all its elements are the same). Had we used any other constant vector, we would have also got the $\mathbf{0}$ -vector.

6. $\mathbf{M} = \mathbf{I} - \frac{1}{n} \mathbf{1} \mathbf{1}^T$

Here, the matrix \mathbf{M} is the matrix we used in the above 4 and 5 exercice. It has dimension $n \times n$, and its content is :

- On the diagonal, all the element are $1 - \frac{1}{n}$
- All the other elements are $-\frac{1}{n}$

We've seen that one of its usage when multiplying a vector is to center it around its mean while preserving its variance.

7. $\mathbf{1}^T \mathbf{x}$

This result in a scalar that is equal to the sum of all the element of \mathbf{x} . $1 \times n$ and $n \times 1 \Rightarrow$ scalar.

8. $\mathbf{1}^T \mathbf{M} \mathbf{x}$

This result in 0. We've established earlier that using this \mathbf{M} matrix result in a vector whose mean is centered around 0. When adding all the elements together (which is what we do when we take the dot product $\mathbf{1}^T \mathbf{M} \mathbf{x}$), we thus get 0.

Quick dimension check : $\mathbf{1}^T \Rightarrow 1 \times n$. $\mathbf{M} \mathbf{x} \Rightarrow n \times 1$. $1 \times n \cdot n \times 1 \Rightarrow$ scalar. $\mathbf{M} \mathbf{x}$ is equal to the question 3 and 4.

9. $\mathbf{M}\mathbf{M}$

Here, we're doing matrix multiplication of \mathbf{M} with itself. However, one of the property of this matrix is that it is idempotent, meaning once it is applied once, applying it again does nothing. As such, $\mathbf{M}\mathbf{M} = \mathbf{M}$. Extrapolating, $\mathbf{M}\mathbf{M}\mathbf{x} = \mathbf{M}\mathbf{x}$.

To understand this, we can think about how, once a vector is projected onto something by the \mathbf{M} matrix, re-applying the matrix does not do anything : the vector is already where it needs to be, centered around its mean. Re-centering it would do nothing : it's already centered.

This idempotency is a property of all projection Matrix, and is proved [here](#)

10. $\mathbf{x}^T\mathbf{x}$

This is the standard formula for taking the dot product of \mathbf{x} with itself. The result is a scalar equal to the sum of all the element squared of \mathbf{x} . This is also equal to the norm squared of \mathbf{x} .

11. $(\mathbf{M})^T(\mathbf{M})$

Here, we're testing exercise 9 again, but this time with a transposed \mathbf{M} matrix. However, \mathbf{M} is symmetric : the element above and below the diagonal are all the same, so transposing it does nothing. Functionnaly speaking : $(\mathbf{M})^T = \mathbf{M}$.

As such, $(\mathbf{M})^T(\mathbf{M}) = \mathbf{M}\mathbf{M} = \mathbf{M}$ still.

12. $\mathbf{x}^T\mathbf{y}$

This is the formula to take the dot product with two vector. Assuming both vector have the same number of element, this result in a scalar that is the sum of the element-wise product of the two vector.

13. $(\mathbf{M}\mathbf{x})^T(\mathbf{M}\mathbf{y})$

In this equation, we're getting close to calculating the covariance between \mathbf{x} and \mathbf{y} . To remember, the equation of the covariance is :

$$\sigma_{xy} = \frac{1}{n-1} \sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y)$$

Here, by taking $\mathbf{M}\mathbf{x}$, we've effectively arrived to $(x_i - \mu_x)$. By then taking the dot product between those two centered vector \mathbf{x} and \mathbf{y} , we're effectively computing the sum that the part $\sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y)$ computes. It is a scalar.

We still need $\frac{1}{n-1}$ to compute the covariance however.

14. $\frac{1}{n-1}(\mathbf{M}\mathbf{x})^T(\mathbf{M}\mathbf{y})$

This is the covariance formula for \mathbf{x} and \mathbf{y} . We've seen in the 13. exercise that $\mathbf{M}\mathbf{x}^T\mathbf{M}\mathbf{y} = \sum_{i=1}^n (x_i - \mu_x)(y_i - \mu_y)$, and that all we needed what the first $\frac{1}{n-1}$ factor at the beginning.

15. $\mathbf{M}\mathbf{X}$

We've established that \mathbf{M} was a nxn matrix, and we know that \mathbf{X} is a nxk matrix. As such, the result of $\mathbf{M}\mathbf{X}$ will be a nxk matrix.

Given that applying $\mathbf{M}\mathbf{x}$ transform \mathbf{x} into its demeaned version, it is nice to know that $\mathbf{M}\mathbf{X}$ is pretty straightforward : it demean every columns of \mathbf{X} by its mean, effectively demeaning the entire matrix.

We can think of it like that.

$$\$ \mathbf{X} = \{ \mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_k \} \$$$

where every \mathbf{x}_i is a column/feature vector. Here, $\mathbf{MX} = \{ \mathbf{Mx}_1, \mathbf{Mx}_2, \dots, \mathbf{Mx}_k \}$. And we know that \mathbf{Mx}_i is simply \mathbf{x}_i demeaned.

16. $(\mathbf{MX})^\top (\mathbf{MX})$

Here, we'll be multiplying the demeaned matrix by itself. We've seen that \mathbf{MX} produces a $n \times k$ matrix with each column demneaed matrix. Transposing the first part of the formula mean making a $k \times n \times n \times k$ matrice multiplication, resulting in a final $k \times k$ matrice.

This is close to resulting in the covariance matrix, but it's once again missing that final $\frac{1}{n-1}$ factor to get there that would normalize it. What we have here is a square matrix where each element is the sum of the deviation between each pair of i, j features, where $i, j = 1 \dots k$.

- The diagonal entries (when $i = j$) represent the sum of squared deviations from the mean for each column i, j .
- The off-diagonal entries (when $i \neq j$) represent the sum of the product of deviations from the mean between each pair of columns i, j .

If we wanted to get the covariance matrix, we'd need to factor the result by $\frac{1}{n-1}$:

$\frac{1}{n-1} (\mathbf{MX})^\top (\mathbf{MX})$. This would result in a symetrical, square, $k \times k$ matrix where the diagonal are 1 (since the pair i, j of row,column are actually the same thing originally before the transpose, and thus have a perfect covariance.), while the other elements would be the covariance between each pair of i, j features.

Add-on : This serie of exercice doesn't directly give us the variance as per the equation that was mentioned at the start of the exercice. If we wanted to compute the variance of \mathbf{x} like in the formula given, but with vector, we could do this :

$$\sigma_x^2 = \frac{1}{n-1} \|\mathbf{Mx}\|^2$$

Ex-2

Part 1 : Definition of linear dependance

The definition of linear dependence for a set of vectors can be written as :

$$\beta_1 \alpha_1 + \dots + \beta_k \alpha_k = 0$$

where the set of vectors α is linearly dependant if there are solutions for the notation to be corret other than by setting all $\beta_i = 0$.

Is $\{0, 0\}$ linearly dependent?

We can multiply the $\{0, 0\}$ vector by any $\beta \in \mathbb{R}$ and still get a correct value. Example :

$$\beta \cdot \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \vec{0}$$

Correct values include :

$$1 \cdot \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \vec{0}, \quad 2 \cdot \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \vec{0}, \quad -1 \cdot \begin{bmatrix} 0 \\ 0 \end{bmatrix} = \vec{0}.$$

Since the set of possible answers β is more than just 0, it is linearly dependent.

Are $\{1, 0\}$ and $\{2, 10\}$ linearly dependent?

Since $\{2, 10\} = 2\{1, 5\}$, they are linearly dependent.

c) Show that $x = (1, 1, 0)$, $y = (0, 1, 2)$ are linearly dependent.

$$\alpha \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix} + \beta \begin{bmatrix} 0 \\ 1 \\ 2 \end{bmatrix} + \gamma \begin{bmatrix} 3 \\ -1 \\ -4 \end{bmatrix} = \vec{0}$$
$$\Rightarrow \begin{bmatrix} 1 & 0 & 3 \\ 1 & 1 & -1 \\ 0 & 1 & -4 \end{bmatrix} \begin{bmatrix} \alpha \\ \beta \\ \gamma \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}$$
$$\Rightarrow \begin{cases} 1\alpha + 0\beta + 3\gamma = 0 \\ 1\alpha + 1\beta - 1\gamma = 0 \\ 0\alpha + 1\beta - 4\gamma = 0 \end{cases} \Rightarrow \begin{cases} \alpha + 3\gamma = 0 \\ \alpha + \beta - \gamma = 0 \\ \beta - 4\gamma = 0 \end{cases}$$

Since we can then plug any γ , there are infinite amount of possibilities for infinitely many equations to hold.

$\begin{pmatrix} 3 \\ 1 \\ -4 \end{pmatrix}$ are linearly dependent

$$\begin{matrix} 3 \\ 1 \\ -4 \end{matrix} = \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -2 \\ 0 & 2 & -4 \end{bmatrix}^{-1} \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{aligned} \alpha &= -3y \\ \beta &= 2y \\ y &= y \end{aligned} \Rightarrow (\alpha, \beta, y) = (-3y, 2y, y)$$

\mathbb{R} value here, we have an
that work to make the

Note : When checking for linear (in)dependence, we do not need to write the $\vec{0}$ target vector inside of the matrix, because it would always remain 0 no matter what. So, we do not need to build an augmented matrix. The process remain the exact same nonetheless.

$$\begin{aligned} \alpha &= \begin{bmatrix} 1 \\ 4 \\ 5 \end{bmatrix} + \beta \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} + y \begin{bmatrix} 2 \\ 4 \\ 4 \end{bmatrix} \\ &\Rightarrow \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} + \beta \begin{bmatrix} 2 \\ 2 \\ 3 \end{bmatrix} + y \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \\ &\quad \text{free variable } \alpha, \beta \end{aligned}$$

$\begin{vmatrix} 1 & 2 & 3 \\ 2 & 2 & 4 \\ 1 & 2 & 2 \end{vmatrix} = 1(2 \cdot 2 - 4 \cdot 2) - 2(2 \cdot 2 - 4 \cdot 1) + 1(2 \cdot 4 - 2 \cdot 2) = 1(-8) - 2(-4) + 1(0) = -8 + 8 = 0$

The vectors $(1, 2, 3)$ and $(2, 2, 4)$ are linearly dependent.

Part 2 : Définition of linear independence

A set of vector α are linearly independent if the only way to make this equation hold is by setting all β_i to 0 :

$$\beta_1\alpha_1 + \cdots + \beta_k\alpha_k = 0$$

If at least one β_i is not 0, and the equation still hold, then the set is linearly dependent.

Ex a) Show that the vectors $(1,2)$ and $(3,4)$ are linearly independent.

For small quantities of small vectors, one easy test is to see if we can obtain one vector by multiplying a set of the others :

$$\beta_1 \begin{bmatrix} 1 \\ 2 \end{bmatrix} = - \begin{bmatrix} 3 \\ 4 \end{bmatrix}$$

(Note : this above notation is just the first equation re-written)

Here, it's obvious that there are no β value that make this equation true. So, the set of vector is linearly independent. This becomes much harder to do with more, bigger vectors. I'll do the exercice b) with a standard method.

Part 2

b) Show that $x = \begin{pmatrix} 3 \\ 4 \end{pmatrix}$, $v = \begin{pmatrix} 5 \\ 2 \end{pmatrix}$ are linearly dependent

$$\begin{bmatrix} 3 & 5 \\ 4 & 2 \end{bmatrix} \leftarrow \text{make pivot: } \frac{1}{3} = \begin{bmatrix} 1 & \cancel{\frac{5}{3}} \\ 4 & 2 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 \\ 4 & 2 \end{bmatrix} \leftarrow \text{make pivot: } \frac{1}{4} = \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{4} \end{bmatrix}$$

$\Rightarrow \begin{bmatrix} 1 & 0 \\ 0 & -\frac{1}{4} \end{bmatrix} \begin{bmatrix} 3 \\ 4 \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \quad \Rightarrow \quad \alpha + \alpha \beta = 0 \quad \text{and} \quad \alpha + \beta = 0 \Rightarrow \alpha = 0, \beta = 0$

independent

$$\begin{pmatrix} \alpha, \beta \\ 0, 0 \end{pmatrix} = \begin{pmatrix} 5/3 \\ 7 \end{pmatrix} - \frac{5}{3} \begin{pmatrix} 1 \\ 1 \end{pmatrix}$$

The matrix would be 0

Part 3

Let $w = (1, 1, 0, 0)$, $x = (1, 0, 1, 0)$, $y = (0, 0, 1, 1)$, and $z = (0, 1, 0, 1)$.

Part a :

P-2 Suite
x-2 Suite

Part 3

Let $w = (1, 1, 0, 0)$, $x = (1, 0, 1, 0)$, $y = (0, 0, 1, 1)$, $z = (0, 1, 0, 1)$

a) $u = (1, 2, 3, 2)$

$$\Rightarrow \left[\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 2 \\ 0 & 1 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 & 2 \end{array} \right] \xrightarrow{-1} \left[\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 2 \end{array} \right] \xrightarrow{-1} \left[\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 & 2 \end{array} \right]$$

$$\left[\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 1 \\ 0 & -1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 & 2 \end{array} \right] \xrightarrow{\cdot 1} \left[\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 & 2 \end{array} \right] \xrightarrow{+1} \left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 2 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 & 2 \end{array} \right] \xrightarrow{\cdot 1} \left[\begin{array}{cccc|c} 1 & 2 & 1 & 1 & 2 \\ 0 & 1 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 3 \\ 0 & 0 & 1 & 1 & 2 \end{array} \right]$$

$$\left[\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 4 \\ 0 & 0 & 1 & 1 & 2 \end{array} \right] \xrightarrow{(-1)} \left[\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 4 \\ 0 & 0 & 1 & 1 & 2 \end{array} \right]$$

$$\left[\begin{array}{cccc|c} 1 & 1 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 & 4 \\ 0 & 0 & 0 & 0 & 2-4 \end{array} \right] \Rightarrow \begin{array}{l} C_1 + C_2 + 0 + 0 = 1 \\ 0 - C_2 + 0 + C_4 = 1 \\ C_3 + C_4 = 4 \\ 0 + 0 + 0 + 0 = 2-4 \end{array} \Rightarrow \begin{array}{l} C_1 = -C_2 = -(C_4 - 1) \\ C_2 = C_4 - 1 \\ C_3 = 4 - C_4 \\ \boxed{d = 4} \end{array}$$

The system is only correct when $d = 4$.

So, when $d = 4$: the system is consistent, it has answer.

When $d \neq 4$, the system is inconsistent

\Rightarrow When a system is inconsistent, it means its product is NOT part of the span

So :

$$\mathbf{u} \in \text{span}\{\mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z}\} \iff a = 4$$

$$\mathbf{u} \notin \text{span}\{\mathbf{w}, \mathbf{x}, \mathbf{y}, \mathbf{z}\} \iff a \neq 4$$

Part b

Ex 2

$$3.2) \quad W = (1.100) \times (1010) \quad y(0011) \quad z(0101)$$

$$\alpha W + x + \gamma y + \delta z = 0$$

$$\Rightarrow \begin{cases} \alpha \cdot 1 + 1 + 2\gamma - 2\delta = 0 \\ \alpha \cdot 1 + \gamma + \delta \cdot 0 + \delta \cdot 1 = 0 \\ \alpha \cdot 0 + 1 + \gamma \cdot 1 + 2\delta = 0 \\ \alpha \cdot 0 + \gamma + \delta \cdot 1 + \delta \cdot 1 = 0 \end{cases} \Rightarrow \begin{cases} \alpha + 1 = 0 \\ \alpha + \delta = 0 \\ 1 + \gamma = 0 \\ \gamma + \delta = 0 \end{cases}$$

1/1

$$\begin{array}{l} \alpha = -1 \\ \delta = -\alpha = 1 \Rightarrow \\ \gamma = -1 \\ \text{dah} \end{array} \quad \{\alpha, \gamma, \delta\} = \{-1, -1, 1\}$$

So :

$$\{\alpha, \gamma, \delta\} = \{-1, -1, 1\}$$

Part c :

TP-2

Ex-2

Part-3 - suite

c) show that $\{w, x, y, z\}$ is linearly dependent by writing z as a product of the others

$$w = (1, 1, 0, 0)$$

$$x = (1, 0, 1, 0)$$

$$y = (0, 0, 1, 1) \Rightarrow c_1 w + c_2 x + c_3 y = z$$

$$z = (0, 1, 0, 1)$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{array} \right] \xrightarrow{\text{R1-R2, R3-R2}} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{array} \right]$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{array} \right] \xrightarrow{\text{R2-R3}} \left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 1 \end{array} \right]$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{array} \right] \xrightarrow{\text{R1-R2}} \left[\begin{array}{ccc|c} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{array} \right]$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 1 & 1 \end{array} \right] \quad c_1 + c_2 = 0 \quad c_1 = -c_2 = 1$$

$$\left[\begin{array}{ccc|c} 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{array} \right] \Rightarrow \begin{array}{l} c_2 = 1 \\ c_3 = 1 \\ 0 = 0 \end{array} \Rightarrow \begin{array}{l} c_1 = -1 \\ c_3 = 1 \\ c_3 = 1 \end{array} = \underbrace{(c_1, c_2, c_3)}_{(1, -1, 1)} = (1, -1, 1)$$

$w - x + y = z$

$$c_1w + c_2x + c_3y = z \iff c_1 = 1, c_2 = -1, c_3 = 1$$

Since there exist an answer other than having all c to 0, the set $\{w, x, y, z\}$ is linearly dependent.

Part 4

4) $u = (\lambda, 1, 0)$ $v = (1, \lambda, 1)$ $w = (0, 1, \lambda)$

$$\alpha u + \beta v + \gamma w = 0$$

$$\Rightarrow \begin{cases} \alpha\lambda + \beta + 0 = 0 & \beta = -\alpha\lambda \\ \alpha + \beta\lambda + \gamma = 0 \\ 0 + \beta + \gamma\lambda = 0 \end{cases} \quad \left\{ \begin{array}{l} \alpha - \alpha\lambda^2 + \gamma = 0 \\ -\alpha\lambda + \gamma\lambda = 0 \\ \lambda(-\alpha + \gamma) = 0 \end{array} \right.$$

$$\alpha - \alpha\lambda^2 + \gamma = 0$$

$$\alpha(1 - \lambda^2) + \gamma = 0 \quad \leftarrow$$

$$\alpha(1 - \lambda^2) + \alpha = 0$$

$$\alpha(1 - \lambda^2 + 1) = 0$$

$$\alpha(2 - \lambda^2) = 0$$

$$2 - \lambda^2 = 0$$

$$\lambda^2 = 2$$

$$\lambda = \sqrt{\pm 2}$$

This is true if $\lambda = 0$

$$-\alpha + \gamma = 0$$

$$\boxed{\gamma = \alpha}$$

$\{u, v, w\}$ are linearly dependent when $\lambda = \{0, \sqrt{-2}, \sqrt{2}\}$

$\{u, v, w\}$ are linearly dependent when $\lambda = \{0, -\sqrt{2}, \sqrt{2}\}$.

Part 5

TP linear algebra

Ex part 5

$$\mathbf{u} = (2 \ 0 \ -1) \quad \mathbf{v} = (3 \ 1 \ 0) \quad \mathbf{w} = (1, -1, c)$$

Set $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ is a basis of \mathbb{R}^3 when $c \neq -2$.

make pivot

$$\left[\begin{array}{ccc|c} 2 & 3 & 1 & \\ 0 & 1 & -1 & \\ -1 & 0 & c & \end{array} \right] \xrightarrow{\text{make pivot}} \left[\begin{array}{ccc|c} 1 & 3/2 & 1/2 & \\ 0 & 1 & -1 & \\ -1 & 0 & c & \end{array} \right]$$

set is linearly independent

since there are only 3 vectors, if they are independent, they will span \mathbb{R}^3

$$\begin{matrix} -1+1 \cdot 1 & 0+1 \cdot 3/2 & c+1 \cdot 1/2 \\ 0 & 3/2 & c+1/2 \end{matrix}$$

$$\left[\begin{array}{ccc|c} 1 & 3/2 & 1/2 & \\ 0 & 1 & -1 & \\ 0 & 3/2 & c+1/2 & \end{array} \right] \xrightarrow{-3/2} \left[\begin{array}{ccc|c} 1 & 3/2 & 1/2 & \\ 0 & 1 & -1 & \\ 0 & 0 & c+2 & \end{array} \right]$$

$$\left[\begin{array}{ccc|c} 1 & 3/2 & 1/2 & 0 \\ 0 & 1 & -1 & 0 \\ 0 & 0 & c+2 & 0 \end{array} \right] \xrightarrow{c+2=0}$$

$$c = -2$$

This is for linear dependence Δ

So, Set $(\mathbf{u}, \mathbf{v}, \mathbf{w})$ is linearly independent, thus a basis for \mathbb{R}^3 when $c \neq -2$

Set $\{\mathbf{u}, \mathbf{v}, \mathbf{w}\}$ is a basis for \mathbb{R}^3 when $c \neq -2$

Part 6

TP Linear
Ex2 Part 6

$$\mathbf{u} = (1 \ 1 \ 1 \ 2) \quad \mathbf{v} = (1 \ 0 \ 1) \quad \mathbf{w} = (1 \ 1 \ 0)$$

$$\begin{bmatrix} 1 & 1 & 1 \\ -1 & 0 & 2 \\ 3 & 1 & c \end{bmatrix} \xrightarrow{\text{R}_1 + R_2, R_3 - 3R_1} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & -2 & c-3 \end{bmatrix}$$

↓

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & -2 & c-3 \end{bmatrix} \xrightarrow{\text{R}_3 + 2R_2} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & c-1 \end{bmatrix}$$

↓

$$\begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & c-1 \end{bmatrix} \xrightarrow{\text{R}_3 + R_2} \begin{bmatrix} 1 & 1 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

$$\left[\begin{array}{ccc|c} 1 & 1 & 1 & 0 \\ 0 & 1 & 1 & 0 \\ 0 & 0 & c-1 & 0 \end{array} \right] \Rightarrow c-1=0$$

$c = -3$ | For linear dependence

So, Set $\{v, u, w\}$ is a basis for \mathbb{R}^3
when $c \neq -3$

Set $\{u, v, w\}$ is a basis for \mathbb{R}^3 when $c \neq -3$

Part 7

TP linear

Ex 2

Part 7

$$x = (1, 1, 1) \quad y = (0, 2, 1) \quad z = (3, 1, 0)$$

1) Check if $\{x, y, z\}$ is a basis of \mathbb{R}^3

$$\alpha x + \beta y + \gamma z = 0$$

$$\begin{bmatrix} 1 & 0 & 3 \\ 1 & 2 & 1 \end{bmatrix} \xrightarrow{\text{R}_2 - R_1} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 2 & 0 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 1 & 0 \end{bmatrix} \sim \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \sim \begin{bmatrix} 1 & -1 & 1 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$\begin{bmatrix} 1 & 0 & 3 \\ 0 & 2 & -2 \\ 0 & 1 & -3 \end{bmatrix} \xrightarrow{\text{swap}} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -3 \\ 0 & 2 & -2 \end{bmatrix} \xrightarrow{-2} \begin{bmatrix} 1 & 0 & 3 \\ 0 & 1 & -3 \\ 0 & 0 & 4 \end{bmatrix} \xrightarrow{4=0} \text{doesnt work}$$

$$\begin{array}{c|c} 1 & 0 & 3 \\ 0 & 1 & -3 \\ 0 & 0 & 4 \end{array} \Rightarrow 4=0 \text{ doesn't work}$$

So $\{x, y, z\}$ is linearly independent.

$$2) W = (3, 1, 1)$$

$$\alpha x + \beta y + \gamma z = W$$

$$\begin{bmatrix} 1 & 0 & 3 & | & 3 \\ 1 & 1 & 1 & | & 1 \\ 1 & 1 & 0 & | & 1 \end{bmatrix} \xrightarrow{-1} \begin{array}{ccc|cc} 1 & -1 & 1 & 2 & 10 \\ 0 & 2 & -2 & 2 & -2 \\ 0 & 1 & -3 & -2 & -2 \end{array} \xrightarrow{2+1} \begin{array}{ccc|cc} 1 & -1 & 1 & 1 & 1 \\ 0 & 2 & -2 & 0 & 10 \\ 0 & 1 & -3 & 1 & -3 \end{array} \xrightarrow{2 \times 2} \begin{array}{ccc|cc} 1 & -1 & 1 & 1 & 1 \\ 0 & 1 & -3 & 0 & 5 \\ 0 & 1 & -3 & 1 & -3 \end{array} \xrightarrow{1+2} \begin{array}{ccc|cc} 1 & 0 & -2 & 1 & 6 \\ 0 & 1 & -3 & 0 & 5 \\ 0 & 1 & -3 & 1 & -3 \end{array} \xrightarrow{1+3} \begin{array}{ccc|cc} 1 & 0 & -2 & 1 & 3 \\ 0 & 1 & -3 & 0 & 5 \\ 0 & 0 & 0 & 1 & 2 \end{array}$$

$$\begin{bmatrix} 1 & 0 & 3 & | & 3 \\ 0 & 2 & -2 & | & -2 \\ 0 & 1 & -3 & | & -2 \end{bmatrix} \xrightarrow{\text{swap}} \begin{bmatrix} 1 & 0 & 3 & | & 3 \\ 0 & 1 & -3 & | & -2 \\ 0 & 2 & -2 & | & -2 \end{bmatrix} \xrightarrow{-2} \begin{array}{ccc|cc} 1 & 0 & 3 & 0 & 1 \\ 0 & 1 & -3 & 2 & -1 \\ 0 & 2 & -2 & 0 & 1 \end{array} \xrightarrow{1+2} \begin{array}{ccc|cc} 1 & 0 & 3 & 0 & 1 \\ 0 & 1 & -3 & 2 & -1 \\ 0 & 0 & 1 & 4 & 2 \end{array} \xrightarrow{1+3} \begin{array}{ccc|cc} 1 & 0 & 3 & 0 & 1 \\ 0 & 1 & -3 & 2 & -1 \\ 0 & 0 & 1 & 4 & 2 \end{array} \xrightarrow{\text{make pivot}} \begin{array}{ccc|cc} 1 & 0 & 3 & 0 & 1 \\ 0 & 1 & -3 & 2 & -1 \\ 0 & 0 & 1 & 1 & 2 \end{array} \xrightarrow{-3+3 \cdot 1} \begin{array}{ccc|cc} 1 & 0 & 3 & 0 & 1 \\ 0 & 1 & 0 & 2 & -1 \\ 0 & 0 & 1 & 1 & 2 \end{array} \xrightarrow{-3+3 \cdot 1} \begin{array}{ccc|cc} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 2 & -1 \\ 0 & 0 & 1 & 1 & 2 \end{array}$$

$$\begin{bmatrix} 1 & 0 & 0 & | & 1.5 \\ 0 & 1 & 0 & | & -0.5 \\ 0 & 0 & 1 & | & 1/2 \end{bmatrix} \xrightarrow{\frac{3}{2}x - \frac{1}{2}y + \frac{1}{2}z = W} \left[\begin{array}{c} W \\ \{x, y, z\} \end{array} \right] = \left(\frac{3}{2}, -\frac{1}{2}, \frac{1}{2} \right)$$

In the basis of $\{x, y, z\}$, w can be expressed as $\left(\frac{3}{2}, -\frac{1}{2}, \frac{1}{2}\right)$

Part 8

First part, converting to orthonormal basis :

TP linear algebra
Ex 2 p. 8

1) Convert to orthonormal basis

$$x = (1, 1, 1) \quad y = (0, 2, 1) \quad z = (3, 1, 0)$$

$$X' = \frac{X}{\|X\|} = \frac{X}{\sqrt{1^2 + 1^2 + 1^2}} = \frac{1}{\sqrt{3}} X = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}$$

only need maximisation

$\checkmark)$ orthogonal $= y - (x'^T y)x'$ scalar

$$= \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} - \left(\begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{3} & 1/\sqrt{3} \end{bmatrix} \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \right) \cdot \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} = \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

normalize:

$$\frac{y}{\|y\|} = \frac{1}{\sqrt{1^2 + 1^2}} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix} = y'$$

$\checkmark)$ orthogonal $= z - (x'^T z)x' - (y'^T z)y'$

$$= z - \frac{4/\sqrt{3}}{\sqrt{2}} x' - (-\sqrt{2}) y'$$

$$= \begin{bmatrix} 3 \\ 1 \\ 0 \end{bmatrix} - \begin{bmatrix} 4/\sqrt{3} \\ 4/\sqrt{3} \\ 4/\sqrt{3} \end{bmatrix} - \begin{bmatrix} 1 \\ -1 \\ 0 \end{bmatrix}$$

$$= \begin{bmatrix} 2/\sqrt{3} \\ 2/\sqrt{3} \\ -4/\sqrt{3} \end{bmatrix}$$

normalize: $\frac{z}{\|z\|} = \frac{z}{2\sqrt{6}/\sqrt{3}} = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{bmatrix}$

So :

$$x' = \begin{bmatrix} \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \\ \frac{1}{\sqrt{3}} \end{bmatrix}, \quad y' = \begin{bmatrix} -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \\ 0 \end{bmatrix}, \quad z' = \begin{bmatrix} \frac{1}{\sqrt{6}} \\ \frac{1}{\sqrt{6}} \\ -\frac{2}{\sqrt{6}} \end{bmatrix}.$$

Second part : find w using the orthonormal basis.

Since we computed the orthonormal basis, we can find the coef to w by simply taking the dot product between w and x' , and do so for the rest for the vectors as well. The product is the coef for x' to multiply it to obtain w .

TP
Ex 2 Part 8 suite

$$x' = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix}, y' = \begin{bmatrix} -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix}, z' = \begin{bmatrix} 1/\sqrt{6} \\ 1/\sqrt{6} \\ -2/\sqrt{6} \end{bmatrix}$$

$$w = \begin{pmatrix} 3 & 1 & 1 \end{pmatrix}$$

Instead of solving $w = a_1 x' + a_2 y' + a_3 z'$
 for the coef, we can find the coeffs like that:

$$a_1 = w^T \cdot x' \\ = (3 \ 1 \ 1) \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} = \underline{\frac{5}{\sqrt{3}}}$$

$$a_2 = w^T \cdot y' \\ = -\sqrt{2}$$

$$a_3 = w^T \cdot z' \\ = \sqrt{2/3}$$

So :

$$w = a'_1 x' + a'_2 y' + a'_3 z' \iff \{ a'_1, a'_2, a'_3 \} = \left\{ \frac{5}{\sqrt{3}}, -\sqrt{2}, \sqrt{\frac{2}{3}} \right\}$$

3 : Convolution

(Note for future me : If you forgot what a convolution is, just check [this](#), it's veryyyyyy good)

Part 1 : 1D signal

1. The function conv1d, takes several arguments: x, kernel, padding, stride, please define each of this terms and how they are used in the convolution.
 - i. x is the signal itself, a timeserie of some sort where every element after the first has some ties to its precedent element, be it simply because it is the one of the day/hour/etc... after, or because the value is based on some operation done on the precedent value. In essence, it's a 1-d vector with n instance.
 - ii. The kernel can be thought of as some sort of function. It takes a vector in input (all the values of the time serie that can fit given the current position of the cursor i and the current center of the kernel) and produce a scalar in output, which, in this case, is the sum of all the element in input divided by 5.
 - iii. The padding is something that can be used to add some element to the base timeserie to avoid the problem that arise when the entire kernel can not be fitted in the timeserie yet. This happen at the start and at the end of the unpadded timeserie, when the elements that would be the window, given the current position of i, are less than kernel_len. Without padding, this may cause trouble with out of bound error, incorrect values, or may cause the output to be smaller than the base signal. It happens in my base conv1d method, where the signal is 1000 elements long and the output result is 996 elements long.
 - iv. The stride is by how much we move the cursor i every time. It has a few effects : first, it changes the size of the result like this : $\text{len(result)} = \text{len(signal)}/\text{stride}$. As the stride augment, the effect that the kernel applies is made more apparent. If too big, it may output a completely unpredictable, useless result, especially if $\text{stride} > \text{kernel_len}$
2. What can you say about the kernel we choose to apply to the time series? If the time series is the evolution of a stock price, then what is the convolution doing?
 - i. We applied a kernel like this : $[1/5, 1/5, 1/5, 1/5, 1/5]$, which will take an input of its same size (5) and take the dot product to produce a scalar. Here, with this kernel, it means that the input values are summed and then divided by 5. In other word, the new, convoluted value is an average of those 5 inputs.
 - ii. If this was the evolution of the stock price, it would mean that the value at the current position of i would be smoothed out based on the values of the 2 next and past days, taking the average values of those 4 days as well as the day at i to produce the new, convoluted value. In more economic terms, this become the moving average of the days around it.
3. Look at the time series before and after the convolution, are the curves the same? If they aren't, explain the effect of the convolution.
 - i. The curves are not the same, but they are still similar. What has been done to them is that they have been smoothed out, they aren't as jagged as before. This is because the kernel we designed takes the values of the elements around a given point and averages them out, meaning that, after convolution, the values are all a little bit closer to each

other, which is what causes this smoothing effect.

Part 2 : 2D convolution.

In 2D convolution, the easiest way to think about it is with an image, as an image is, after all, just a 2D matrix of pixel (as long as it's in grayscale). Then, our kernel is also a matrix, but its function is still the same. For a given point i, j , the kernel capture all the points around it that fit the kernel (once again, think of the kernel as being centered around the point, and checking which values fits in it), multiplie the values of those points by the value of their element wise pair in the kernel, sum all of this up, and that sum is the value of the new, convoluted i, j value.

First kernel : Edge detection

The first kernel :

$$\begin{bmatrix} -1 & -1 & -1 \\ -1 & 8 & -1 \\ -1 & -1 & -1 \end{bmatrix}$$

is an edge detection kernel. It is designed to emphasize areas in an image where there are rapid changes in intensity, which often correspond to edges.

This kernel works by giving a strong, positive weight to the center pixel (8) and subtracting the values of its surrounding neighbors with weights of -1. The result of this operation reflects the difference between the center pixel's grayscale value (intensity) and the average values of its neighboring pixels. When the kernel is applied to an area with little to no change in grayscale values (for example, a flat surface or uniform color), the contributions of the neighbors (nearly) cancel out the strong central value, leading to a value close to 0, which is black-ish in a greyscale image. When the kernel is applied to an area with a sharp change in grayscale values (for example, a boundary between a dark and a light region), the center pixel's value will differ significantly from its neighbors. This difference produces a large positive value, which is more white.

Second kernel : Sharpening

The second kernel :

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & -1 & 5 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

is a sharpening kernel. It starts off simialr to an edge detection kernel, with a big positive center

value surrounded by low negative value. However, the key difference is in the sum of the kernel : if sum = 0, then the kernel is an edge detection kernel. If sum > 0, it's a sharpening kernel.

In effect, this kernel works the same way as an edge detection kernel. When in a smooth, uniform region, the value of the center pixel gets negated by its neighbors, but because the center is always valued a little bit more than the sum of the neighbors (since the sum > 0), the final value end up staying close to the original value, which preserve the value mostly close to the original value.

When covering an area with a change in values like an edge, the kernel works the exact same way as an edge detection kernel.

Third Kernel : Gaussian blur

The kernel

$$\frac{1}{256} \cdot \begin{bmatrix} 1 & 4 & 6 & 4 & 1 \\ 4 & 16 & 24 & 16 & 4 \\ 6 & 24 & 36 & 24 & 6 \\ 4 & 16 & 24 & 16 & 4 \\ 1 & 4 & 6 & 4 & 1 \end{bmatrix}$$

is a blurring kernel, more specifically Gaussian blur. All blurring kernel follow the same overall logic : the value of the center pixel become an average of all the pixels surrounding it. At its most basic, all pixel are valued equal with a kernel that looks like this :

$$\frac{1}{9} \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix}$$

A gaussian blur kernel is a variant of this that gives more importance to the pixels closer to the center and slowly lower the importance for the pixels that are further away, following a Gaussian curve. This is done symmetrically. The result on the image is a more natural feeling of blur.

The normalization factor ($\frac{1}{256}$) is there to assure that the overall sum of the kernel is 1. This help in making sure than any result from the blurring effect keeps the same general tone, and doesn't suddenly become a lot brighter or darker.

4 : Rotation of a 2D dataset.

The matrix is sourced from [here](#).

The intuition to better understand is a video [here](#)

Explain what you see in the above figure, what does the rotation matrix do to the dataset? Does the distance change between the points in the dataset after a rotation ? If the distance between the points changes, give the formula to compute the new distance.

What does the matrix do to the dataset ?

Unsurprisingly, what a rotation matrix do to a dataset is pretty self-explanatory : it rotates it by a defined degree (WOAW).

To understand how it works, let's start with only a single (unit) vector $[1, 0]$, on the unit circle. We can imagine rotating it θ degree counterclockwise, keeping its tail on the origine while its head move θ along the unit circle. Because its length (or hypothenuse) stay the same (1), we can still compute its coordinates using trigonometry :

(Note to futur me : if this isn't clear, check the vid, it's way more visual, I can't do (good) pictures here.)

$$\cos \theta = \frac{x}{\text{hypotenuse}} \rightarrow \cos \theta = x$$

similarly, for the y value of this vector :

$$\sin \theta = \frac{y}{\text{hypotenuse}} \rightarrow \sin \theta = y$$

So

$$\text{rotated}[1, 0] = [\cos \theta, \sin \theta]$$

If we now take a second vector on the unit vector, this time $[0, 1]$, and rotate it in the same direction by θ as well, its coordinate can once again be obtained via trigonometry :

$$\sin \theta = \frac{x}{\text{hypotenuse}} \iff \sin \theta = x$$

Though, here, in this case, the value of x is negative. So, it's actually $-\sin$

Still, we also have :

$$\cos \theta = \frac{y}{\text{hypotenuse}} \rightarrow \cos \theta = y$$

So :

$$rotated[0, 1] = [-\sin\theta, \cos\theta]$$

With this setup, we have the rotation function for the two unit vector of a 2D matrix :

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix}$$

which is simply the two vector we found earlier, put vertically. And since any vector $[x, y]$ can be thought of as a scaling of the two unit vector, this formula applies to any 2D vector.

From there, we can apply this to a dataset, thinking of each instance n as its own vector $[x, y]$ to apply the transformation to :

$$\begin{bmatrix} \cos \theta & -\sin \theta \\ \sin \theta & \cos \theta \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$$

Does the transformation conserve the distance between points ?

As we've seen in our example, the length (or norm) of a rotated vector was kept through the transformation. Because we apply the transformation to every n instance through the dataset uniformly, all their own length stay the same, and they keep pointing in the same, rotated direction so, the overall position of each points compared to each other stay the same.

In the notebook, the cell before the last one shows how this intuition is true.