Linear Algebra
Week 3

G-07

10 X 2024

1 CR Decomposition

A matrix decomposition is breaking a matrix in several matrices, which usually each give us information about the decomposed matrix. CR decomposition is the first of many matrix decompositions you are going to encounter during the course. From the lecture notes:

Theorem 2.23 (CR decomposition) *Let A be an $m \times n$ matrix of rank r. Let C be the $m \times r$ submatrix of A containing the independent columns. Then there exists a unique $r \times n$ matrix R such that:*

$$A = CR$$

Let's see why this is true.

- The column space of A: $\mathbf{C}$(A) = $\mathbf{C}$(C) : Column space of C. This is because we have all linearly independent columns of A in C.
- But first point means that we can write the columns of A as a linear combination of the columns of C. This is because the columns of A themselves are also in the column space of A!
- With this point of view, in R at i-th column, we store how much of which column of C we need to generate the i-th column of A.

$$\underbrace{\begin{bmatrix} 1 & 2 & 0 & 3 \\ 2 & 4 & 1 & 4 \\ 3 & 6 & 2 & 5 \end{bmatrix}}_{A,\ 3\times 4} = \underbrace{\begin{bmatrix} 1 & 0 \\ 2 & 1 \\ 3 & 2 \end{bmatrix}}_{C,\ 3\times 2} \underbrace{\begin{bmatrix} 1 & 2 & 0 & 3 \\ 0 & 0 & 1 & -2 \end{bmatrix}}_{R,\ 2\times 4}.$$

The 4^{th} column of R tells us that we need 3 times the 1^{st} column and -2 times the 2^{nd} column of C, in order to generate the 4^{th} column of A.

Question 3.1 Calculate the matrices C and R for $A = \begin{bmatrix} 3 & 9 & -3 & 12 \\ 1 & 10 & -8 & 11 \\ 8 & 7 & 9 & 15 \end{bmatrix}$

2 Linear Transformations

Linear Transformations are one of the fundamental concepts of linear algebra. Under the hood matrices are tools to represent linear transformations. At this point we should consider matrices as functions.

2.1 Matrices as Functions

Matrices are functions! They are defined as follows in the lecture notes:

Definition 2.25 (Matrix as function) *Let A be an $m \times n$ matrix. $T_A : \mathbb{R}^n \mapsto \mathbb{R}^m$ is the function where $\boldsymbol{x} \in \mathbb{R}^n$ and $A\boldsymbol{x} \in \mathbb{R}^m$ defined by:*

$$T_A(\mathbf{x}) = A\mathbf{x}$$

This shows that every matrix defines a linear transformation.

There is an important set of observations about matrices as functions, that show that these transformations are linear (below they are going to be defined as axioms of linear transformations).

Observation 2.26. *Let A be an $m \times n$ matrix, $\mathbf{x}, \mathbf{y} \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$. Then*

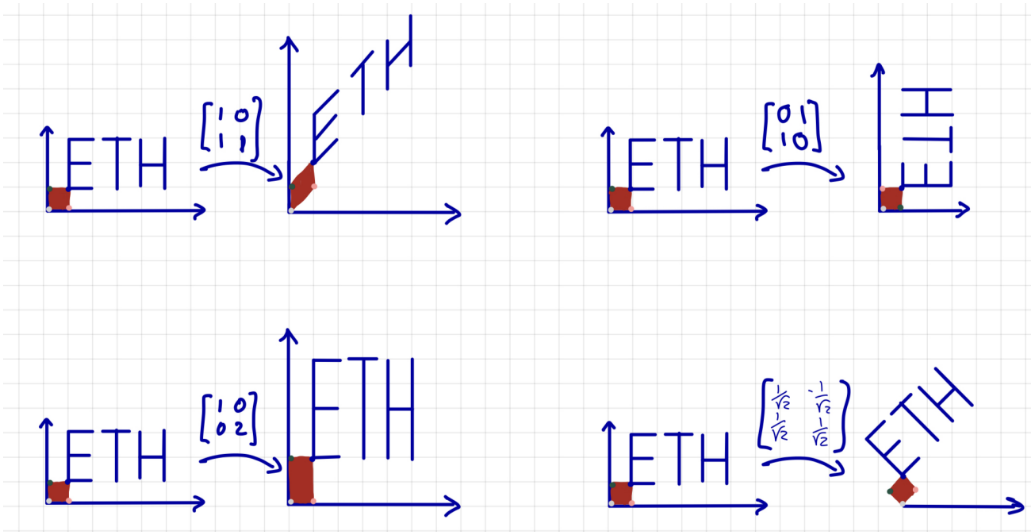
(i) $T_A(\mathbf{x} + \mathbf{y}) = T_A(\mathbf{x}) + T_A(\mathbf{y})$ *and*

(ii) $T_A(\lambda\mathbf{x}) = \lambda T_A(\mathbf{x})$.

By combining (i) and (ii), we also get $T_A(\lambda\mathbf{x} + \mu\mathbf{y}) = \lambda T_A(\mathbf{x}) + \mu T_A(\mathbf{y})$.

To understand in what ways a matrix transforms a vector we are specifically interested in the standard unit vectors $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$. To intuitively sum up why, it is because we can express everything in a space $\mathbb{R}^n$ using the set of standard unit vectors. If we know what A does to this set of vectors, we can generalize the effect of A to all vectors in $\mathbb{R}^n$.

This is demonstrated below. If you know what happens to the unit square (little red guy), then you can generalize that effect to any other shape or *e.g.* to any set of vectors.



Example from HS23

2.2 Linear Transformations

We already have the intuition on linear transformations since they are equivalent to the function a matrix represents. Here is a formal definition from the lecture notes:

Definition 2.27 (Linear transformation) *Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a function from $\mathbb{R}^n$ to $\mathbb{R}^m$. T is called a linear transformation if the following two statements hold for all $x, y \in \mathbb{R}^n$ and all $\lambda \in \mathbb{R}$.*

(i) $T(x + y) = T(x) + T(y)$ and
(ii) $T(\lambda x) = \lambda T(x)$.

By combining (i) and (ii), it then also holds that $T(\lambda x + \mu y) = \lambda T(x) + \mu T(y)$.

(i) and (ii) are called the *axioms of a linear transformation*. In words this means that it is not important with which order we apply the operations: We can first add two vectors and then apply T, or first apply T to the vectors individually and then add the results. They are both equal. Same principle also applies for scaling *i.e.* multiplying with a scalar. Hence the equation holds: $T(\lambda x + \mu y) = \lambda T(x) + \mu T(y)$

An important lemma that formalizes this for more than two vectors, is the following which comes with a proof by induction that I recommend reading from the lecture notes.

Lemma 2.28. *Let $T : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a linear transformation, let $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_\ell \in \mathbb{R}^n$ and $\lambda_1, \lambda_2, \ldots, \lambda_\ell \in \mathbb{R}$. Then*

$$T\left(\sum_{j=1}^{\ell} \lambda_j \mathbf{x}_j\right) = \sum_{j=1}^{\ell} \lambda_j T(\mathbf{x}_j).$$

In words, the function value of a linear combination is the linear combination of the function values.

0 is always mapped to **0** by a linear transformation: $T(\mathbf{0}) = \mathbf{0}$

2.3 Matrix of a linear Transformation

We have seen that every matrix A defines a linear transformation T_A above (Def 2.25). This theorem states the other direction: Every linear transformation $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ (Definition 2.27) is of the form $T = T_A$ for a unique $m \times n$ matrix A.

Theorem 2.29. Let $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ be a linear transformation. There exists a unique $m \times n$ matrix A such that $T = T_A$.

Proof: For this theorem the only candidate for a suitable matrix A is the following:

$$A = \begin{bmatrix} | & | & & | \\ v_1 & v_2 & \dots & v_n \\ | & | & & | \end{bmatrix} = \begin{bmatrix} | & | & & | \\ A \cdot \mathbf{e}_1 & A \cdot \mathbf{e}_2 & \dots & A \cdot \mathbf{e}_n \\ | & | & & | \end{bmatrix}$$

$$\overset{Def 2.25}{=} \begin{bmatrix} | & | & & | \\ T_A(\mathbf{e}_1) & T_A(\mathbf{e}_2) & \dots & T_A(\mathbf{e}_n) \\ | & | & & | \end{bmatrix}$$

Since $T_A(\mathbf{e}_i)$ is the i-th column of A. After convincing us to this fact, e can proceed with:

$$T_A(\mathbf{x}) = A\mathbf{x} = \sum_{j=1}^{n} x_j T(\mathbf{e}_j) = T\left(\sum_{j=1}^{n} x_j \mathbf{e}_j\right) = T(\mathbf{x}).$$

The first equality is just the definition 2.25. The second equality takes a few steps at the same time. $A\mathbf{x}$ is the linear combination of the columns of A as we defined it, which is $\sum_{j=1}^{n} x_j v_j$ where v_j are the columns of A. Using that A is our candidate matrix from above, we can write $T(\mathbf{e}_j)$ instead of v_j as the j-th column of A. This brings us to the third equation which is just an implementation of Lemma 2.28. Last but not least, the fourth equation is justified because $\sum_{j=1}^{n} x_j \mathbf{e}_j$ is equivalent to multiplying the vector $\mathbf{x}$ with the identity matrix.

□

2.4 Matrix Multiplication and Linear Transformation

We can combine linear transformations and represent the composition as matrix multiplication. Let $T_A : \mathbb{R}^n \mapsto \mathbb{R}^a$ and $T_B : \mathbb{R}^a \mapsto \mathbb{R}^m$ be two linear transformations. Then their composition $T : \mathbb{R}^n \mapsto \mathbb{R}^m$ is:

$$T(\mathbf{x}) = T_A(T_B(\mathbf{b}))$$

The composition is again a linear transformation itself and it is represented with the matrix $C = AB$.

Here is the formalized version of the lemma and its proof from the lecture notes:

Lemma 2.30. *Let* $T_A : \mathbb{R}^n \to \mathbb{R}^a$ *and* $T_B : \mathbb{R}^b \to \mathbb{R}^n$ *be two linear transformations. Then*

$$T_A(T_B(\mathbf{x})) = T_{AB}(\mathbf{x}) \text{ for all } \mathbf{x} \in \mathbb{R}^b.$$

Proof. This is a one-liner:

$$T_A(T_B(\mathbf{x})) = T_A(B\mathbf{x}) = A(B\mathbf{x}) = (AB)\mathbf{x} = T_{AB}(\mathbf{x}).$$

Here, we have used Definition 2.25 for T_A, T_B, T_{AB} as well as associativity of matrix multiplication (Lemma 2.22), where we treat $\mathbf{x}$ as a $b \times 1$ matrix. □

Question 3.2 What are the dimensions of $\mathbf{x}, T(\mathbf{x}), T_B(\mathbf{x}), T_A(\mathbf{x})$ and $T_A(T_B(\mathbf{x}))$?

2.5 Image and Kernel

Here are the definitions for Image and Kernel of a matrix (from lecture notes). We are going to examine them closer in a couple of weeks but the keywords already come up in assignments and sometimes in the lecture.

For every linear transformation, there are two important sets of vectors.

Definition 2.31 (Kernel and image). *Let $T : \mathbb{R}^n \to \mathbb{R}^m$ be a linear transformation. The set*

$$\mathbf{Ker}(T) := \{\mathbf{x} \in \mathbb{R}^n : T(\mathbf{x}) = \mathbf{0}\} \subseteq \mathbb{R}^n$$

is the kernel *of T. The set*

$$\mathbf{Im}(T) := \{T(\mathbf{x}) : \mathbf{x} \in \mathbb{R}^n\} \subseteq \mathbb{R}^m$$

is the image *of T.*

3 Linear Equations

We are going to handle equations of type $A\mathbf{x} = \mathbf{b}$ in this chapter. This week was introduction to this part, so here is a brief representation of how to think of matrices as linear equations in m equations and n variables:

Definition 3.1 (System of linear equations). *A* system of linear equations *in m equations and n variables $x_1, x_2, \ldots, x_n$ is of the form*

$$\begin{array}{rcl} a_{11}x_1 + a_{12}x_2 + \cdots + a_{1n}x_n & = & b_1 \\ a_{21}x_1 + a_{22}x_2 + \cdots + a_{2n}x_n & = & b_2 \\ & \vdots & \\ a_{m1}x_1 + a_{m2}x_2 + \cdots + a_{mn}x_n & = & b_m, \end{array}$$

where the a_{ij} and b_i stand for known real numbers, and the x_i stand for unknown real numbers that we want to compute such that they satisfy all the equations. In matrix-vector form, this can be written as

$$A\mathbf{x} = \mathbf{b} : \quad \underbrace{\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{bmatrix}}_{A,\ m \times n} \underbrace{\begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}}_{\mathbf{x} \in \mathbb{R}^n} = \underbrace{\begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}}_{\mathbf{b} \in \mathbb{R}^m}.$$

Observation 3.2. *Let A be an $m \times n$ matrix. The columns of A are linearly independent if and only if the system $A\mathbf{x} = \mathbf{0}$ has a unique solution, $\mathbf{x} = \mathbf{0}$.*

Proof. By Definition 2.4 of matrix-vector multiplication, the solutions are precisely the vectors of scalars that express $\mathbf{0}$ as a linear combination of the columns. A unique solution means that $\mathbf{0}$ can only be written as a trivial linear combination of the colums. By Lemma 1.19, this is equivalent to the columns of A being linearly independent. □

4 Hints

1. In Class.

2. Bonus! No hints.

3. First two subtasks are about trigonometric knowledge. For c) use your finding from b).

4. You might want to look at the formal definition of a triangular matrix again. Write A and B in column and row notation respectively and multiply them. How can we represent the elements of the matrix AB using the columns of A and the rows of B? What do we know about them? for b) use a).

5. Think of an example. Consider the effect of this matrix on the standard unit vectors. By the way, $sin(\frac{\pi}{4}) = cos(\frac{\pi}{4}) = \frac{1}{\sqrt{2}}$ For b) again think about the unit vectors and what is supposed to happen to them in order to realize the rotation around the line spanned by $\mathbf{v} = \begin{bmatrix}1 & 1 & 0\end{bmatrix}^\top$.

6. Treat vectors $\begin{bmatrix}1 & 0\end{bmatrix}^\top$ and $\begin{bmatrix}1 & 1\end{bmatrix}^\top$ as variables and calculate $T(\begin{bmatrix}1 \\ 0\end{bmatrix})$. How can you use the standard unit vectors to come up with a general formula for the matrix $T(\begin{bmatrix}x \\ y\end{bmatrix})$?

7. No hints.

mkilic

♠