Orthogonality and Applications of Linear Algebra

Lecture 6

Zhang Ce

Department of Electrical and Electronic Engineering
Southern University of Science and Technology

2022.11.8

Table of Contents

Table of Contents

Inner Product (Dot Product)

Actually, you have learnt that in your senior high school...

If I give you 2 vectors $\begin{bmatrix} 1 \\ 2 \end{bmatrix}, \begin{bmatrix} 3 \\ 1 \end{bmatrix}$, how to compute its inner products?

$$\begin{bmatrix} 1 \\ 2 \end{bmatrix} \cdot \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 1 \times 3 + 2 \times 1 = 5$$

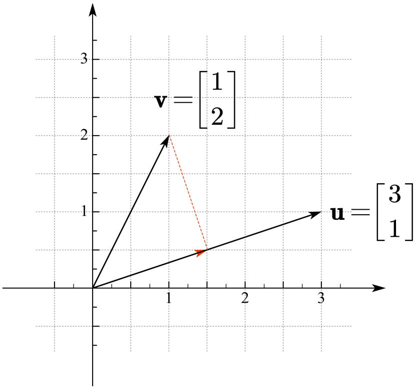
Recall matrix multiplications, which rule for matrix multiplication is similar to this?

$$\begin{bmatrix} 1 & 2 \end{bmatrix} \begin{bmatrix} 3 \\ 1 \end{bmatrix} = 3 + 2 = 5$$

Therefore, for 2 vectors u and v, the inner product is $u^T v$.

Take a deeper look, the result 5 shows us... Think in geometrical way!

Inner Product (Dot Product)



In my perspective, their inner product reflects their relevance. If the inner product can take the maximum or minimum value, then the vectors are dependent. While if the inner product is greater than zero, they point to the same direction, if it equals to zero, they have nothing to do with each other (perpendicular).

Orthogonal Vectors and Subspaces

Which value of inner product can let you realize that the vectors are perpendicular (orthogonal)?

$$u^T v = 0$$

Then we can further extend this definition to subspaces, if all the vectors in subspace A are orthogonal to all the vectors in subspace B, then subspaces A and B are orthogonal.

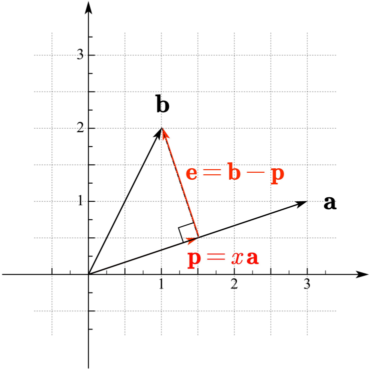
One good question to ask: Is the blackboard plane perpendicular to the ground? But are they orthogonal?

Well, two orthogonal subspaces can only have the origin in common, and that is from the restriction of subspaces!

Table of Contents

Projections

Let's start from geometrical view. We want to find projection $\mathbf{p}$.



What can we get from that 90 degrees? Can we get an expression of $\mathbf{p}$?

$$\mathbf{a}^T\mathbf{e} = \mathbf{a}^T(\mathbf{b} - x\mathbf{a}) = 0 \Rightarrow x = \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}} \Rightarrow \mathbf{p} = \mathbf{a}\frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}$$

What if the length of vector $\mathbf{a}$ doubles? How about $\mathbf{b}$?

Projection Matrix

$$\mathbf{a}^T\mathbf{e} = \mathbf{a}^T(\mathbf{b} - x\mathbf{a}) = 0 \Rightarrow x = \frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}} \Rightarrow \mathbf{p} = \mathbf{a}\frac{\mathbf{a}^T\mathbf{b}}{\mathbf{a}^T\mathbf{a}}$$

Can you give a projection matrix that can project every vector b onto the line where a lies in?

$$\mathbf{p} = P\mathbf{b} \Rightarrow P = \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}$$

That is the matrix for projection onto line a. Use your knowledge of previous knowledge, $\mathbf{a}\mathbf{a}^T$ is a scalar, vector or matrix? What about $\mathbf{a}^T\mathbf{a}$?

Think in linear transformation, what is the rank of that matrix P, that is to ask the dimension of the output space (the column space).

Definitely 1, I can also get the column space is the line containing vector $\mathbf{a}$.

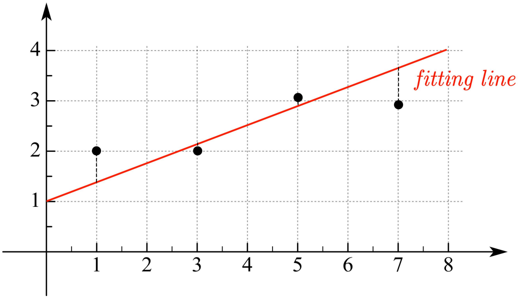
Some properties for projection matrix P?

Table of Contents

Motivation

In Chapter 2, when we meet a linear equation system $Ax = b$, it may have no solutions. But we can find a solution that can minimize the error! In practice, if we have more equations than the variables, we may want to find the best approximation solution.

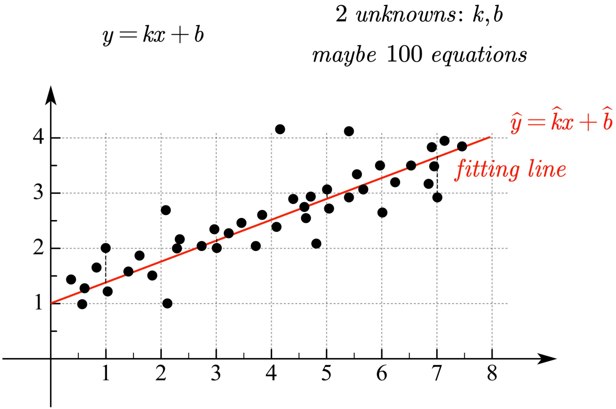
Take fitting as an example:



We can't find a line to cover all the points, but we can find the line that minimizes the error.

Motivation

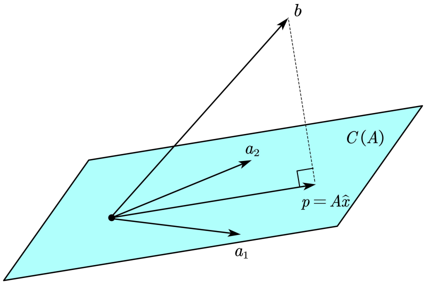
Take fitting as an example:



How to find the predicted line? The answer is: least-squares.

Least-Squares

The core is: find the solution for $Ax = p$ (A is the matrix with a_1, a_2 in columns) where p is the projection on $C(A)$, which can definitely gives a unique solution.



$$a_1{}^T(b - A\hat{x}) = a_2{}^T(b - A\hat{x}) = 0$$

The error $e = b - p = b - A\hat{x}$ is in the nullspace of A, while the projection $p = A\hat{x}$ is in the column space of A!

Least-Squares

Well, we can simplify that equation...

$$a_1{}^T(b - A\hat{x}) = a_2{}^T(b - A\hat{x}) = 0$$

$$\begin{bmatrix} a_1{}^T \\ a_2{}^T \end{bmatrix}(b - A\hat{x}) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

$$A^T(b - A\hat{x}) = 0$$

$$A^T A\hat{x} = A^T b$$

Now, can you tell me why $A^T A\hat{x} = A^T b$ must be consistent? The geometric reason behind: $Ax = p$ must have solutions.

How to prove? Hint: $N(A^T A) = N(A)$ introduced on Review Class.

Finally, we can get:

$$\hat{x} = (A^T A)^{-1} A^T b$$

That is the least-squares solution, indicating how to combine the columns in A can we get the nearest vector to b.

Projection Matrices

The projection vector in the column space, which is also the nearest vector to b in $C(A)$:

$$p = A\hat{x} = A\left(A^TA\right)^{-1}A^Tb = Pb$$

Now we can have projection matrices onto column space of A:

$$P = A\left(A^TA\right)^{-1}A^T$$

Can I take the inverse and simplify the expression further?

$$P = A\left(A^TA\right)^{-1}A^T = AA^{-1}(A^T)^{-1}A^T = I\,???$$

A is not square matrix, A^{-1} doesn't exist.

Recall that we have introduced projection matrices onto lines:

$$P = \frac{\mathbf{a}\mathbf{a}^T}{\mathbf{a}^T\mathbf{a}}$$

It is just a special case (1-D case) for projection matrices.

Table of Contents

Introduction to This Part

Linear Algebra is definitely a mathematical subject, but it isn't in the Advanced Mathematics (Calculus) course. That is because linear algebra is the most (not one of the most) useful course in university! The thought and proposed method in linear algebra is useful in all the subject, including chemistry, biology, electrical engineering and computer science...

We are now studying to use a tool, just like the four elementary computations you learnt in your childhood. With these tools, the complicated concept becomes easier.

You are not supposed to understand all of the knowledge in this part, but, try to feel the beauty of linear algebra.

Linear Algebra in Chemistry

How can linear algebra be used in chemistry?

Recall your chemistry knowledge in your senior high school, try to balance the following chemical reaction equation:

$$NO_2 + H_2O \rightarrow HNO_3 + NO$$

You may use these knowledge:

- Mass conservation
- Element conservation

Here comes Bob, he knows nothing about chemistry, he even doesn't know what N and O represent for. But he is really good at linear algebra, he do the problem by this way:

Firstly, he set all the coefficients with variables x_1, x_2, x_3, x_4.

$$x_1 NO_2 + x_2 H_2 O \rightarrow x_3 HNO_3 + x_4 NO$$

- For element N: $x_1 = x_3 + x_4$.
- For element O: $2x_1 + x_2 = 3x_3 + x_4$.
- For element H: $2x_2 = x_3$.

$$A = \begin{bmatrix} 1 & 0 & -1 & -1 \\ 2 & 1 & -3 & -1 \\ 0 & 2 & -1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 2 & -1 & 0 \end{bmatrix} \rightarrow \begin{bmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & -2 \end{bmatrix} = U$$

$$U = \begin{bmatrix} 1 & 0 & -1 & -1 \\ 0 & 1 & -1 & 1 \\ 0 & 0 & 1 & -2 \end{bmatrix} \to \begin{bmatrix} 1 & 0 & 0 & -3 \\ 0 & 1 & 0 & -1 \\ 0 & 0 & 1 & -2 \end{bmatrix} = R$$

He solve the $Ax = 0$ system by nullspace matrix, and he gets that:

$$x = \begin{bmatrix} 3 \\ 1 \\ 2 \\ 1 \end{bmatrix}$$

The result is

$$3NO_2 + H_2O = 2HNO_3 + NO$$

If you multiply all the coefficients by a constant, the chemical reaction equation is still balanced, but not the simplest.

Not excited? Try this one.

$$P_4 + CuSO_4 + H_2O \to Cu_3P + H_2SO_4 + H_3PO_4$$

- For element P: $4x_1 = x_4 + x_6$.
- For element Cu: $x_2 = 3x_4$.
- For element S: $x_2 = x_5$.
- For element O: $4x_2 + x_3 = 4x_5 + 4x_6$.
- For element H: $2x_3 = 2x_5 + 3x_6$.

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

Guess the rank of the matrix, give explanations also.

$$\begin{bmatrix} 4 & 0 & 0 & -1 & 0 & -1 \\ 0 & 1 & 0 & -3 & 0 & 0 \\ 0 & 1 & 0 & 0 & -1 & 0 \\ 0 & 4 & 1 & 0 & -4 & -4 \\ 0 & 0 & 2 & 0 & -2 & -3 \end{bmatrix} \rightarrow \begin{bmatrix} 4 & 0 & 0 & -1 & 0 & -1 \\ 0 & 1 & 0 & -3 & 0 & 0 \\ 0 & 0 & 1 & 12 & -4 & -4 \\ 0 & 0 & 0 & 3 & -1 & 0 \\ 0 & 0 & 0 & 0 & -2 & 5 \end{bmatrix}$$

Free variable x_6 set to 2, the solution is:

$$x = \begin{bmatrix} 11/12 & 5 & 8 & 5/3 & 5 & 2 \end{bmatrix}^T$$

Multiply a constant 12, the solution becomes:

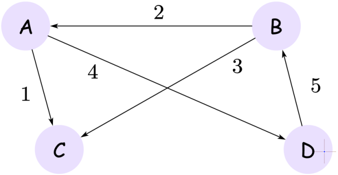
$$x = \begin{bmatrix} 11 & 60 & 96 & 20 & 60 & 24 \end{bmatrix}^T$$

So, the final chemical reaction equation is:

$$11P_4 + 60CuSO_4 + 96H_2O = 20Cu_3P + 60H_2SO_4 + 24H_3PO_4$$

Linear Algebra in Graphs

Every single graph can be represented by a matrix.



Nodes

$$
\begin{array}{c c}
\begin{array}{cccc} A & B & C & D \end{array} & \\
\begin{bmatrix} -1 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 \\ 0 & -1 & 1 & 0 \\ -1 & 0 & 0 & 1 \\ 0 & 1 & 0 & -1 \end{bmatrix} & \begin{array}{c} 1 \\ 2 \\ 3 \\ 4 \\ 5 \end{array}
\end{array}
$$

Edges

Eliminate it to RREF. Find the nullspace solution.

$$\# \text{ of nodes } -1 = r$$

Dependent rows represent there exists a loop.

$$\# \text{ of loops} = dim(N(A^T))$$

We can get the Euler's Formula (holds for all graphs):

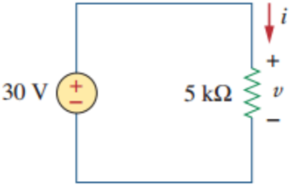
$$dim(N(A^T)) = m - r \rightarrow \# \text{ of loops} = \# \text{ of edges - (\# of nodes - 1)}$$

Linear Algebra in Circuit Principles

Three important theorems in circuits:

- Ohm's Law: $I = U/R$ for resistors
- Kirchhoff's Current Law: $\sum I = 0$ for each node
- Kirchhoff's Voltage Law: $\sum V = 0$ for each mesh

Ohm's Law:



$$i = v/R = \frac{30V}{5k\Omega} = 6mA$$

Kirchhoff's Current Law (KCL):

By KCL,

At node A, $I_1 + I_2 = 2A$

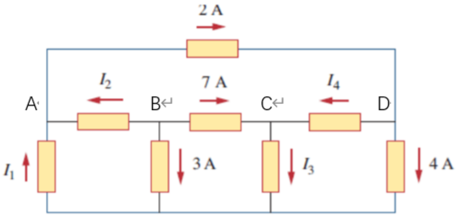
At node B, $0 = I_2 + 10A$

At node C, $I_4 + 7A = I_3$

At node D, $2A = I_4 + 4A$

By solving 4 equations, we get

$$I_1 = 12A,\ I_2 = -10A,\ I_3 = 5A,\ I_4 = -2A$$



Kirchhoff's Voltage Law (KVL):

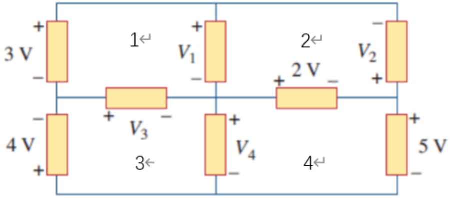
By KVL,

In loop 1, $V_1 - V_3 - 3V = 0$

In loop 2, $V_1 + 2V + V_2 = 0$

In loop 3, $V_3 + V_4 + 4V = 0$

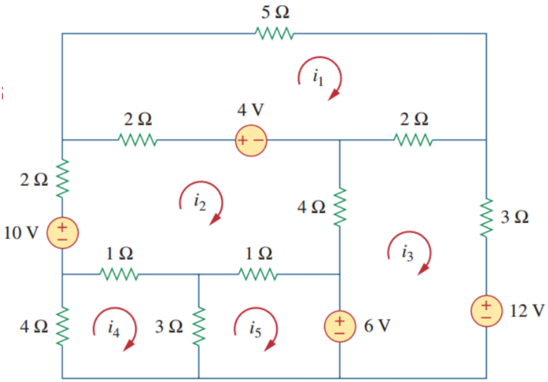
In loop 4, $V_4 - 5V - 2V = 0$



By solving 4 equations, we get

$V_1 = -8V$, $V_2 = 6V$, $V_3 = -11V$, $V_4 = 7V$

Seems to be easy... But for complicated circuit, can you figure out the whole state of it?



Our method is: find the mesh current i_1, i_2, i_3, i_4, i_5.

The essence: Solve $Ax = b$ linear system!

$$\begin{bmatrix} 9 & -2 & -2 & 0 & 0 \\ -2 & 10 & -4 & -1 & -1 \\ -2 & -4 & 9 & 0 & 0 \\ 0 & -1 & 0 & 8 & -3 \\ 0 & -1 & 0 & -3 & 4 \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \\ i_3 \\ i_4 \\ i_5 \end{bmatrix} = \begin{bmatrix} 4 \\ 6 \\ -6 \\ 0 \\ -6 \end{bmatrix}$$

Every row is a KVL equation for a mesh.

If there is no voltage source, then the equation system becomes $Ax = 0$, the mesh currents are definitely all zero (that is to say the matrix is of full rank because all the meshes are independent, no repeat meshes occur)!

The result will be

$$i_1 = 0.283A, i_2 = 0.211A, i_3 = -0.488A, i_4 = -0.718A, i_5 = -1.986A$$

Computer Vision is a great subject in artificial intelligence. Let's have a brief introduction about how our images are like in the computer view.

In a computer, how does it read or store an image?

256×256 *picture*

$$\longrightarrow$$

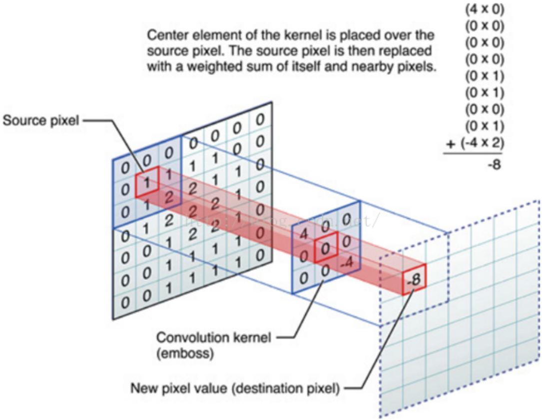
$$\begin{bmatrix} 137 & 136 & 222 & 133 & 245 & 221 & 135 & 168 \\ 135 & 231 & 56 & 43 & 211 & 213 & 135 & 147 \\ 146 & 45 & 112 & 211 & 98 & 21 & 33 & 21 \\ 145 & 23 & 198 & \ddots & \cdots & \cdots & \cdots & 22 \\ 143 & 111 & 186 & \vdots & \ddots & & & 33 \\ 111 & 124 & 175 & \vdots & & \ddots & & 44 \\ 197 & 211 & 111 & \vdots & & & \ddots & 55 \\ 211 & 222 & 111 & 254 & 211 & 11 & 32 & 66 \end{bmatrix}$$

256×256 *matrix*

0 *for black*, 255 *for white*

It is a natural matrix! If a picture is colorful, then we may have 3 256×256 matrices for R, G, B respectively.

A new operation: convolution.



Now, let's see what will happen for different convolution kernel.

- Identity Filter:

$$* \begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \longrightarrow$$

- Sharpness Filter (enhance contents):

$$* \begin{bmatrix} -1 & -1 & -1 \\ -1 & 9 & -1 \\ -1 & -1 & -1 \end{bmatrix} \longrightarrow$$

- Another Sharpness Filter (enhance edges):

$$* \begin{bmatrix} 1 & 1 & 1 \\ 1 & -7 & 1 \\ 1 & 1 & 1 \end{bmatrix} \longrightarrow$$

- Edge Detection Filter:

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

Linear Algebra in Computer Vision

- Average Box Filter:

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

- Gauss Smoothing Filter:

$$* \ \frac{1}{16}\begin{bmatrix} 1 & 2 & 1 \\ 2 & 4 & 2 \\ 1 & 2 & 1 \end{bmatrix} \longrightarrow$$

Linear Algebra in Computer Vision

Gauss Smoothing Filter for many times:

Original *2 times Smoothing* *5 times Smoothing* *10 times Smoothing*

Every time when you use the mobile phone, if you try to pull down the notification badge, Gauss Smoothing Filter is applied. That is the magic from Linear Algebra!