Gram - Schmidt

Let $\{\underline{e}_1, \underline{e}_2\}$ be a basis of $\mathbb{R}^2$

Define
$$\underline{q}_1 = \underline{e}_1 / \|\underline{e}_1\| = \underline{e}_1 / (\underline{e}_1^T \underline{e}_1)^{1/2}$$

and
$$\tilde{\underline{q}}_2 = \underline{e}_2 - \underline{q}_1^T \underline{e}_2 \, \underline{q}_1$$

$$\underline{q}_2 = \tilde{\underline{q}}_2 / \|\tilde{\underline{q}}_2\|$$

$\left\{ \begin{array}{l} \underline{q}_1 \text{ and } \underline{q}_2 \text{ are linear combinations} \\ \text{of } \underline{e}_1 \text{ and } \underline{e}_2 \end{array} \right.$

similarly

$$\underline{e}_1 = \underline{q}_1 \|\underline{e}_1\| \quad \text{and} \quad \underline{e}_2 = \underline{q}_2 \|\tilde{\underline{q}}_2\| + \underline{q}_1^T \underline{e}_2 \, \underline{q}_1$$

so

$\left\{ \begin{array}{l} \underline{e}_1 \text{ and } \underline{e}_2 \text{ are linear combinations of} \\ \underline{q}_1 \text{ and } \underline{q}_2 \end{array} \right.$

i.e. $\boxed{\underline{q}_1 \text{ and } \underline{q}_2 \text{ span } \mathbb{R}^2}$

$$\underline{q}_1^T \tilde{\underline{q}}_2 = \underline{q}_1^T \underline{e}_2 - \underline{q}_1^T \underline{e}_2 \, \underline{q}_1^T \underline{q}_1 = 0$$

so $\underline{q}_1$ and q_2 are orthogonal and have norm one

$$\boxed{\text{$\underline{q}_1$ and $\underline{q}_2$ form an orthonormal basis for $\mathbb{R}^2$}}$$

they are linearly independent since

$$\alpha_1 \underline{q}_1 + \alpha_2 \underline{q}_2 = 0 \Rightarrow \underline{q}_1^T (\alpha_1 \underline{q}_1 + \alpha_2 \underline{q}_2) = 0$$

$$\Rightarrow \alpha_1 = 0$$

and similarly for α_2

if we had a third basis vector $\underline{e}_3$ (say in $\mathbb{R}^3$) the we would define

$$\left[\begin{array}{l} \tilde{\underline{q}}_3 = \underline{e}_3 - (\underline{q}_1^T \underline{e}_3) \underline{q}_1 - (\underline{q}_2^T \underline{e}_3) \underline{q}_2 \\ \underline{q}_3 = \tilde{\underline{q}}_3 / \| \tilde{\underline{q}}_3 \| \end{array}\right.$$

to get an orthonormal basis

$\{\underline{q}_1, \underline{q}_2, \underline{q}_3\}$

In general (in $\mathbb{R}^n$) be continue recursively until

$$\boxed{\begin{aligned} \tilde{\underline{q}}_m &= \underline{e}_m - \sum_{j=1}^{m-1} (\underline{q}_j^T \underline{e}_m)\, \underline{q}_j \\ \underline{q}_m &= \tilde{\underline{q}}_m / \|\hat{\underline{q}}_m\| \end{aligned}}$$

Let V be a k dimensional subspace of $\mathbb{R}^n$ and $\{\underline{e}_1 \cdots \underline{e}_k\}$ a basis for V.

Define

$$V^\perp = \{\underline{x} \in \mathbb{R}^n : \underline{x} \cdot \underline{y} = 0 \ \forall \underline{y} \in V\}$$

1st observation: any $\underline{z} \in \mathbb{R}^n$ can be written as $\underline{z} = \underline{x} + \underline{y}$, with $\underline{x} \in V^\perp$ $\underline{y} \in V$.

Pf

based on $\{\underline{e}_1 \cdots \underline{e}_k\}$ use Gram-Schmidt to obtain an orthonormal basis for V. call this $\{\underline{q}_1 \cdots \underline{q}_k\}$. Then define

$$\underline{y} = \sum_{j=1}^{k} (\underline{z} \cdot \underline{q}_j)\, \underline{q}_j \in V$$

$$\underline{x} = \underline{z} - \underline{y} = \underline{z} - \sum_{j=1}^{k} (\underline{z} \cdot \underline{q}_j)\, \underline{q}_j$$ is then in $V^\perp$

(why ??)

and $\underline{z} = \underline{x} + \underline{y}$

$\square$

Call the dimension of $V^\perp$ $\underline{\underline{\underline{\ell}}}$ and let $\{\underline{p}_1 \cdots \underline{p}_\ell\}$ be an orthonormal basis of $V^\perp$ (How did we get that??)

$$\boxed{\text{Then } \{\underline{p}_1 \cdots \underline{p}_\ell\ \underline{q}_1 \cdots \underline{q}_k\} \text{ forms an orthonormal basis for } \mathbb{R}^n}$$

- this set of vectors is clearly linearly independent (why?)
- it spans, since for any $\underline{z} \in \mathbb{R}^n$

$$\underset{\substack{\cap \\ V^\perp}}{\underline{z} = \underline{x}} + \underset{\substack{\cap \\ V}}{\underline{y}} = \sum_{j=1}^{\ell} \alpha_j \underline{p}_j + \sum_{j=1}^{k} \beta_j \underline{q}_j$$

$\square$

$$\boxed{\begin{array}{l}\text{In other words} \\ \dim V + \dim V^\perp = k + \ell = n\end{array}}$$

Let A be an $m \times n$ matrix

We define

$$\mathcal{N}(A) = \{\underline{x} \in \mathbb{R}^n : A\underline{x} = 0\} \quad \left(\begin{matrix}\text{The}\\ \text{Nullspace}\end{matrix}\right)$$

$$\mathcal{R}(A) = \{\underline{y} \in \mathbb{R}^m : \underline{y} = A\underline{x} \text{ for some } \underline{x} \in \mathbb{R}^n\}$$

(The Range)

Similarly $\mathcal{N}(A^T)$ and $\mathcal{R}(A^T)$.

- We note that $A\underline{x} \cdot \underline{y} = \sum_{i,j} A_{ij} x_j y_i$

$$= \underline{x} \cdot A^T \underline{y}$$

- Therefore $\underline{x} \in \mathcal{N}(A) \Leftrightarrow A\underline{x} = 0$

$\Leftrightarrow A\underline{x} \cdot \underline{y} = 0 \ \forall \underline{y} \in \mathbb{R}^m \Leftrightarrow \underline{x} \cdot A^T \underline{y} = 0$

$\forall \underline{y} \in \mathbb{R}^m \Leftrightarrow \underline{x} \in (\mathcal{R}(A^T))^\perp$

or
$$\boxed{\mathcal{N}(A) = (\mathcal{R}(A^T))^\perp}$$

- $\dim\left(\mathcal{R}(A^T)\right)^{\perp} + \dim\left(\mathcal{R}(A^T)\right) = n$

  (why?)

- Thus

$$\dim\left(\mathcal{N}(A)\right) + \dim\left(\mathcal{R}(A)\right)$$
$$= \dim\left(\mathcal{N}(A)\right) + \dim\left(\mathcal{R}(A^T)\right) \qquad \text{why??}$$
$$= \dim\left(\mathcal{R}(A^T)^{\perp}\right) + \dim\left(\mathcal{R}(A^T)\right) = n$$

or

$$\boxed{\underset{\substack{\uparrow \\ \text{nullity}}}{\dim}\left(\mathcal{N}(A)\right) + \underset{\substack{\uparrow \\ \text{rank}}}{\dim}\left(\mathcal{R}(A)\right) = n}$$

when A is $m \times n$!

A is an $\underline{n \times n}$ matrix (square)

We say that $\underline{x}$ is an eigen vector with eigenvalue λ if $\boxed{\underline{x} \neq 0}$

and

$$\boxed{A\underline{x} = \lambda \underline{x}}$$

suppose there exist a basis of n eigenvectors $\{\underline{x}_1 \cdots \underline{x}_n\}$

$$A\underline{x}_j = \lambda_j \underline{x}_j$$

Then

$$A\left(\sum_{j=1}^{n} \alpha_j \underline{x}_j\right) = \sum_{j=1}^{n} \alpha_j A \underline{x}_j$$

$$= \sum_{j=1}^{n} \lambda_j \alpha_j \underline{x}_j$$

in other words the matrix representing A in the basis $\{\underline{x}_1 \cdots \underline{x}_n\}$, i.e., If we change basis to $\{\underline{x}_1 \cdots \underline{x}_n\}$, the transformed matrix is

$$\begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix} = \Lambda$$

or there exists a change of basis matrix $B = [\underline{x}_1 \, \underline{x}_2 \cdots \underline{x}_n]$ so that

$$\boxed{\begin{aligned} B^{-1} A B &= \Lambda \\ &\text{or} \\ AB &= B\Lambda \end{aligned}}$$

The latter simply says

$$[A\underline{x}_1 \, A\underline{x}_2 \cdots A\underline{x}_n] = [\lambda_1 \underline{x}_1 \, \lambda_2 \underline{x}_2 \cdots \lambda_n \underline{x}_n]$$