

Projection with Bases

Note Title

LECTURE 09

★ Projection with an Orthonormal Basis

In general, for a projector $P \in \mathbb{R}^{m \times m}$
 $\dim(\text{range}(P)) = n \leq m$

So, let's define the matrix

$$\hat{Q} := [q_1 \cdots q_n] \in \mathbb{R}^{m \times n}$$

signifies the **reduced** form of Q

where $\{q_1, \dots, q_n\}$ forms an O.N.B.
of $\text{range}(P)$ as in the proof
of the previous theorem.

$$\text{Then, } P = \hat{Q} \hat{Q}^T \quad \left(= Q \Lambda Q^T \right. \\ \left. \text{in the previous thm.} \right)$$

Recall $\forall v \in \mathbb{R}^m, \exists w \in \mathbb{R}^m$ (residual)

$$\text{s.t. } v = w + \sum_{i=1}^n (q_i q_i^T) v$$

$\uparrow \quad \quad \quad \uparrow$
 $\perp \quad \quad \quad \perp$

Hence, the mapping $v \mapsto \sum_{i=1}^n (q_i q_i^T) v = y$

is an orthogonal projection onto $\text{range}(\hat{Q})$

$$\begin{array}{c} 1 \\ \boxed{y} \\ m \end{array} = \begin{array}{c} n \\ \boxed{\hat{Q}} \\ m \end{array} \begin{array}{c} m \\ \boxed{\hat{Q}^T} \\ n \end{array} \begin{array}{c} 1 \\ \boxed{v} \\ m \end{array}$$

as $n \rightarrow m, y \rightarrow v$ since $\hat{Q} \hat{Q}^T \rightarrow I$

LECTURE 09

Note: The complementary proj. to an orth. proj. $P = \hat{Q}\hat{Q}^T$ is also an orth. proj.

why? $I - P = I - \hat{Q}\hat{Q}^T$ is also symmetric! //

Also note the following special case:
The rank-one orth. proj. with a unit vector $\xi \in \mathbb{R}^m$

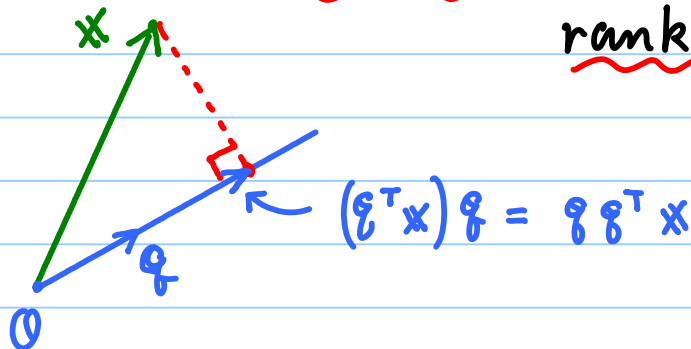
$$P_\xi := \xi\xi^T \in \mathbb{R}^{m \times m}$$

↳ a special rank 1 matrix

Its complementary proj. is

$$P_{\perp \xi} := I - P_\xi =: P_\xi^\perp$$

rank $m-1$ orth. proj.



For a general vector $a \in \mathbb{R}^m$ with $a \neq 0$, $\|a\| \neq 1$, the orth. proj. onto $\text{span}\{a\}$ becomes

$$P_a := \frac{aa^T}{a^T a} \rightarrow P_{\perp a} = I - \frac{aa^T}{a^T a}$$

why? Set $\xi = a/\|a\|$ then it's easy to show. //

LECTURE 09

★ Projection with an Arbitrary Basis

Let $\{a_1, \dots, a_n\} \subset \mathbb{R}^m$ be a set of linearly independent vectors.

Set $A = [a_1 \dots a_n] \in \mathbb{R}^{m \times n}$

assume $m \geq n$.

Then what is the ortho. proj. onto $\text{range}(A) = \text{span}\{a_1, \dots, a_n\}$?

Let $P_A \in \mathbb{R}^{m \times m}$ be such an ortho. proj.

Then $\text{range}(P_A) = \text{range}(A)$

and $\text{null}(P_A) = \text{null}(A^T)$.

Now take any $v \in \mathbb{R}^m$. Then

$\exists v_1 \in \text{range}(A)$, $\exists v_2 \in \text{null}(A^T)$.

s.t. $v = v_1 + v_2$.

$\exists x \in \mathbb{R}^n$ s.t. $v_1 = Ax$ $\perp$ v_2

Hence $a_j \perp v_2 = v - v_1 = v - Ax$
for $1 \leq j \leq n$

$$\Leftrightarrow a_j^T (v - Ax) = 0 \quad 1 \leq j \leq n$$

$$\Leftrightarrow A^T (v - Ax) = 0$$

$$\Leftrightarrow A^T A x = A^T v \quad (\text{normal eqn. !!})$$

HW03 Prob. 6 { Since A is full rank (because $\{a_1, \dots, a_n\}$ are lin. indep.),
 $(A^T A)^{-1}$ exists, i.e., $x = (A^T A)^{-1} A^T v$

Hence $P_A = A (A^T A)^{-1} A^T$

LECTURE 09

We recover the previous projectors by setting

$$\begin{cases} A = \hat{Q} \Rightarrow \hat{Q} (\hat{Q}^T \hat{Q})^{-1} \hat{Q}^T = \hat{Q} \hat{Q}^T \\ A = [\underline{a}] \Rightarrow \underline{a} (\underline{a}^T \underline{a})^{-1} \underline{a}^T = \frac{\underline{a} \underline{a}^T}{\underline{a}^T \underline{a}} \end{cases}$$

$\uparrow$ $m \times 1$ matrix $\hookrightarrow$ a scalar

Ex. $A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix}$

(1) Compute the orthogonal projector P_A onto range(A)

(2) Let $\underline{v} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$. Then compute

its projection onto range(A) using P_A obtained in (1).

Sol. (1) $P_A = A(A^T A)^{-1} A^T$

If $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$ then $A^T A = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 1 \end{bmatrix}$

$A^{-1} = \frac{1}{ad-bc} \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$ $(A^T A)^{-1} = \frac{1}{2} \begin{bmatrix} 1 & 0 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix}$

So, $P_A = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix}$

(2) $P_A \underline{v} = \begin{bmatrix} \frac{1}{2} & 0 & \frac{1}{2} \\ 0 & 1 & 0 \\ \frac{1}{2} & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 2 \\ 2 \\ 2 \end{bmatrix} //$

QR Decomposition

LECTURE 09

* Reduced QR Factorization

Let $A = [a_1 \dots a_n] \in \mathbb{R}^{m \times n}$

Denote $\text{span}\{a_1, \dots, a_k\}$ by
 $\langle a_1, \dots, a_k \rangle$ for simplicity.

Then $\langle a_1 \rangle \subset \langle a_1, a_2 \rangle \subset \langle a_1, a_2, a_3 \rangle$
 $\subset \dots \subset \langle a_1, \dots, a_n \rangle$
Successive column spaces.

QR factorization (or decomposition)

$\therefore$ a method to successively
construct a sequence of
orthonormal vectors $q_1, q_2, \dots$, s.t.
 $\langle q_1, \dots, q_j \rangle = \langle a_1, \dots, a_j \rangle$
 $j = 1, 2, \dots, n$

If this is the case, then each a_j
can be expanded by $\{q_1, \dots, q_j\}$

$$\text{Say, } \begin{cases} a_1 = r_{11} q_1 \\ a_2 = r_{12} q_1 + r_{22} q_2 \\ \vdots \\ a_n = r_{1n} q_1 + r_{2n} q_2 + \dots + r_{nn} q_n \end{cases}$$

$$\Rightarrow A = [q_1 \ q_2 \ \dots \ q_n] \begin{bmatrix} r_{11} & r_{12} & \dots & r_{1n} \\ & r_{22} & \dots & r_{2n} \\ & & \ddots & \vdots \\ & & & r_{nn} \end{bmatrix}$$

$= \hat{Q} \hat{R}$

Reduced QR factorization of A upper triangular

LECTURE 09

★ Full QR Factorization

$$\begin{array}{c}
 \begin{array}{ccc}
 & n & n \quad m-n \\
 m & \boxed{A} = m & \boxed{\hat{Q} \quad \quad} \quad \boxed{\begin{array}{c} n \\ \text{---} \\ n \end{array}} \hat{R} \\
 & & \text{Append } m-n \text{ O.N. vectors to} \\
 & & \text{col's of } \hat{Q} \rightarrow Q \\
 & & \text{Append } m-n \text{ 0's} \\
 & & \text{to rows of } \hat{R} \rightarrow R
 \end{array}
 \end{array}$$

$A = \underbrace{\quad}_Q \underbrace{\quad}_R$

Note: $\xi_j \perp \text{range}(A)$ for $j > n$.
 i.e., $\langle \xi_1, \dots, \xi_n \rangle = \text{range}(A)$
 $\langle \xi_{n+1}, \dots, \xi_m \rangle = \text{range}(A)^\perp$
 $= \text{null}(A^T)$

★ The Classical Gram-Schmidt Orthogonalization as QR

You must be familiar with the classical GS procedure.

Given $\{a_1, \dots, a_n\}$, $a_j \in \mathbb{R}^m$, $1 \leq j \leq n$, construct an orthonormal set $\{\xi_1, \dots, \xi_n\}$, $\xi_j \in \mathbb{R}^m$, $1 \leq j \leq n$ as follows:

$$\xi_1 = \frac{a_1}{r_{11}}, \quad r_{11} = \|a_1\|$$

$$\xi_2 = \frac{a_2 - r_{12}\xi_1}{r_{22}}, \quad \begin{array}{l} r_{12} = \xi_1^T a_2 \\ r_{22} = \|a_2 - r_{12}\xi_1\| \end{array}$$

LECTURE 09

$$\begin{aligned} \xi_3 &= \frac{a_3 - r_{13}\xi_1 - r_{23}\xi_2}{r_{33}}, & r_{i3} &= \xi_i^T a_3, \quad i=1,2. \\ & \vdots & & \vdots \\ \xi_n &= \frac{a_n - \sum_{i=1}^{n-1} r_{in}\xi_i}{r_{nn}}, & r_{nn} &= \|a_n - \sum_{i=1}^{n-1} r_{in}\xi_i\| \end{aligned}$$

So, in general,

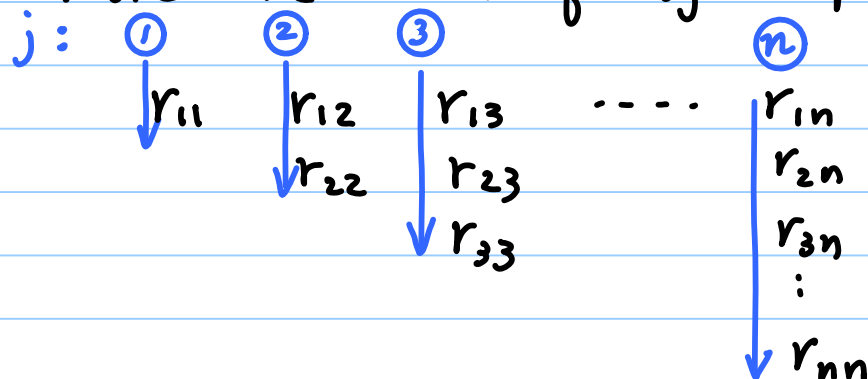
$$\begin{cases} r_{ij} = \xi_i^T a_j & i \neq j \\ r_{jj} = \|a_j - \sum_{i=1}^{j-1} r_{ij}\xi_i\| \end{cases}$$

Algorithm (The classical Gram-Schmidt)

for $j = 1:n$ (CGS)

$$\left\{ \begin{array}{l} v_j = a_j \\ \text{for } i = 1:j-1 \\ \quad \left\{ \begin{array}{l} r_{ij} = \xi_i^T a_j \\ v_j = v_j - r_{ij}\xi_i \end{array} \right. \\ r_{jj} = \|v_j\| \\ \xi_j = v_j / r_{jj} \end{array} \right. \quad \leftarrow \text{error accumulates here}$$

Note the order of r_{ij} computation



LECTURE 09

Unfortunately, this version is numerically unstable.

Ex. $A = \begin{bmatrix} 1 & 1 & 1 \\ \varepsilon & 0 & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & \varepsilon \end{bmatrix}$. ε : small
s.t. ε^2 can
be ignored.

Apply CGS.

$$\begin{cases} r_{11} = \|a_1\| = \sqrt{1^2 + \varepsilon^2 + 0^2 + 0^2} \approx 1. \\ \end{cases}$$

$$\begin{cases} \xi_1 = \frac{a_1}{r_{11}} = a_1. \\ \end{cases}$$

$$\begin{cases} r_{12} = \xi_1^T a_2 = \begin{bmatrix} 1 & \varepsilon & 0 & 0 \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ \varepsilon \\ 0 \end{bmatrix} = 1. \\ \end{cases}$$

$$\begin{cases} \xi_2 = \frac{a_2 - r_{12}\xi_1}{r_{22}} = \frac{1}{r_{22}} \left(\begin{bmatrix} 1 \\ 0 \\ \varepsilon \\ 0 \end{bmatrix} - 1 \cdot \begin{bmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{bmatrix} \right) = \frac{1}{r_{22}} \begin{bmatrix} 0 \\ -\varepsilon \\ \varepsilon \\ 0 \end{bmatrix} \\ \end{cases}$$

$$\begin{cases} r_{22} = \|a_2 - r_{12}\xi_1\| = \varepsilon \sqrt{0 + (-1)^2 + 1^2 + 0^2} = \sqrt{2}\varepsilon \\ \end{cases}$$

$$\begin{cases} \xi_2 = \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \\ 0 \end{bmatrix} \\ \end{cases}$$

$$\begin{cases} r_{13} = \xi_1^T a_3 = 1, \quad r_{23} = \xi_2^T a_3 = 0. \\ \end{cases}$$

$$\begin{cases} \xi_3 = \frac{a_3 - r_{13}\xi_1 - r_{23}\xi_2}{r_{33}} = \frac{1}{r_{33}} \left(\begin{bmatrix} 1 \\ 0 \\ 0 \\ \varepsilon \end{bmatrix} - \begin{bmatrix} 1 \\ \varepsilon \\ 0 \\ 0 \end{bmatrix} \right) \\ \end{cases}$$

$$\begin{cases} = \frac{1}{r_{33}} \begin{bmatrix} 0 \\ -\varepsilon \\ 0 \\ \varepsilon \end{bmatrix} = \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 0 \\ 1/\sqrt{2} \end{bmatrix}, \quad \text{with } r_{33} = \sqrt{2}\varepsilon. \\ \end{cases}$$

LECTURE 09

Hence

$$\hat{Q} = \begin{bmatrix} 1 & 0 & 0 \\ \varepsilon & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix}, \quad \hat{R} = \begin{bmatrix} 1 & 1 & 1 \\ 0 & \sqrt{2}\varepsilon & 0 \\ 0 & 0 & \sqrt{2}\varepsilon \end{bmatrix}$$

Let's check these results.

$$\hat{Q} \hat{R} = \begin{bmatrix} 1 & 1 & 1 \\ \varepsilon & 0 & 0 \\ 0 & \varepsilon & 0 \\ 0 & 0 & \varepsilon \end{bmatrix} = A$$

So, looks OK.

But, how about the orthogonality of $\hat{Q}$?

$$\begin{aligned} \hat{Q}^T \hat{Q} &= \begin{bmatrix} 1 & \varepsilon & 0 & 0 \\ 0 & -\frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} & 0 \\ 0 & -\frac{1}{\sqrt{2}} & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \begin{bmatrix} 1 & 0 & 0 \\ \varepsilon & -\frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ 0 & \frac{1}{\sqrt{2}} & 0 \\ 0 & 0 & \frac{1}{\sqrt{2}} \end{bmatrix} \\ &= \begin{bmatrix} 1+\varepsilon^2 & -\frac{\varepsilon}{\sqrt{2}} & -\frac{\varepsilon}{\sqrt{2}} \\ -\frac{\varepsilon}{\sqrt{2}} & 1 & \frac{1}{2} \\ -\frac{\varepsilon}{\sqrt{2}} & \frac{1}{2} & 1 \end{bmatrix} \\ &\approx \begin{bmatrix} 1 & -\frac{\varepsilon}{\sqrt{2}} & -\frac{\varepsilon}{\sqrt{2}} \\ -\frac{\varepsilon}{\sqrt{2}} & 1 & \frac{1}{2} \\ -\frac{\varepsilon}{\sqrt{2}} & \frac{1}{2} & 1 \end{bmatrix} \\ &\neq I_{3 \times 3} \end{aligned}$$

This is called "loss of orthogonality".