Note Title

QR factorization

Quick review of projection matrices.

Defn: A projection matrix (or a projection $P: V \to V$) is a matrix P if $P^2 = P$.

(Note: if $T: V \to V$ is a projection operator then $[T]$ is a projection matrix)

Properties:

① If $v \in$ range(P), $Pv = v$ { if $v \in$ range P, then $\exists x$ s.t. $Px = v$, $Pv = P^2(x) = Px = v$. }.

② If $v \notin$ range(P), then $P(Pv - v) = P^2 v - Pv = 0$
$\Rightarrow Pv - v \in$ null(P).

③ If P is a projection then $I - P$ is also a projection.

④ range$(I-P) =$ null(P)

⑤ range$(P) =$ null$(I-P)$.

⑥ range$(P) \cap$ null$(P) = \{0\}$.

So P separates the underlying vector space into 2 complementary subspaces i.e. $V =$ range $P \oplus$ null P.

In general, range P need not be orthogonal to null P.

Defn: an orthogonal projection P is one for which range $P \perp$ null P.

Theorem: A projection matrix P is an orthogonal projection $\iff P = P^*$.

Proof: If $P = P^*$, let $Px \in$ range P, $z \in$ null P ($\because (I-P)y$)
$\qquad\qquad\qquad\qquad$ $\overset{\shortparallel}{\text{range}(I-P)}$. $\quad$ z''

Consider $\langle Px, (I-P)y \rangle = ((I-P)y)^* Px$

$= y^* (I-P)^* Px \overset{P=P^*}{=} y^*(P - P^2)x = 0.$

$= y^*(I^* - P^*)Px = y^*(P - P^2)x$

Suppose P is an orthogonal projection.

Let range $P = \langle q_1, \dots, q_n \rangle$, null $P = \langle q_{n+1}, \dots, q_m \rangle$

$\therefore \; P\underline{q_j} = q_j \;\; \forall\, 1 \le j \le n \;\; \& \;\; P\underline{q_j} = 0 \;\; \forall\, n+1 \le j \le m.$

$Q^* = \begin{bmatrix} q_1^* \\ \vdots \\ q_n^* \\ q_m^* \end{bmatrix}$

Let $Q = \left[q_1 \middle| \cdots \middle| q_n \middle| q_{n+1} \middle| \cdots \middle| q_m \right]$

$$PQ = \left[q_1 \middle| \cdots \middle| q_n \middle| 0 \middle| \cdots \middle| 0 \right]$$

$$Q^* P Q = \begin{bmatrix} 1 & & & & & \\ & 1 & & & & \\ & & \ddots & & & \\ & & & 1 & & \\ & & & & 0 & \\ & & & & & \ddots \; 0 \end{bmatrix} \quad \text{(Check)}$$

$$= \Sigma$$

$P = Q \Sigma Q^*$ is the SVD of P.

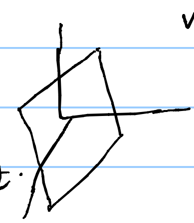
$\therefore \; P^* = (Q \Sigma Q^*)^* = Q \Sigma^* Q^* = Q \Sigma Q^* = P. \quad \square.$

<u>Important</u> <u>note</u>: Orthogonal projections are not orthogonal matrices!!

Projection : $P : V \to V \quad \dim V = m.$

$\dim(\text{range } P) < m. \quad$ (i.e. not full rank)

i.e. P is always rank-deficient.



Can it happen that a projection is full rank?

i.e. range $P = V$.

If range$(P) = V$, then $Pv = v \;\; \forall v \in V$.

$\Rightarrow P = I$.

So for any projection, its matrix is not full rank, so cannot be orthogonal matrix.

Projections $P^2 = P$

- range $P \not\perp$ null P
- range $P \perp$ null P $\to$ orthogonal projection. $P^* = P$

Not orthogonal matrices!

rank k projection : $\dim(\text{range}) = k$.

Special case: rank 1 projections : Let q be a unit vector. (let $\dim V = n$)

Consider $W = \langle q \rangle$, this is a 1-dim'l subspace of V.

What is the projection that maps V onto W?

$$P_q = qq^* \quad \rightarrow \text{this is a rank 1 matrix}$$

Check- $\left[\begin{array}{l} \cdot\ P_q^2 = P_q \\ \cdot\ P_q \text{ maps any } v \in V \text{ onto } W. \end{array} \right.$

The complement $\underset{P_{\perp q}}{I - P_q} = I - qq^*$ is a rank $(m-1)$ matrix.

For an arbitrary vector a, $\quad P_a = \dfrac{aa^*}{a^*a}$

& $\quad P_{\perp a} = I - \dfrac{aa^*}{a^*a}$

$\{a_1, \dots, a_j\}$. $\quad q_i^* a_j = \delta_{ij}$

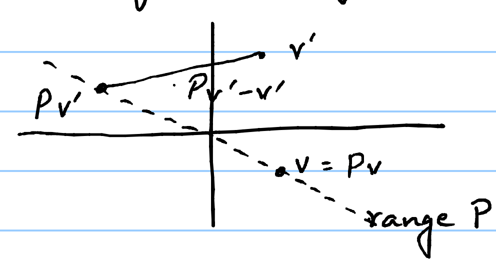
$A = [a_1 | \cdots | a_j]$.

$P_A = AA^*$

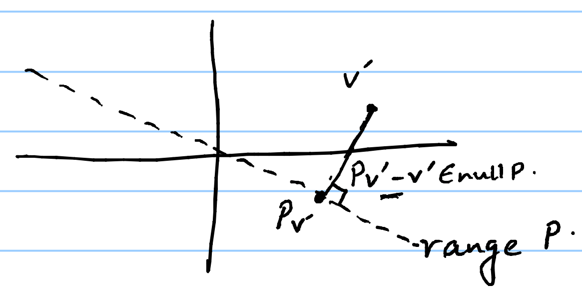
$P_{\perp A} = I - AA^*$

Geometrically (schematic)

- for a non-orthogonal projector P,



- for an orthogonal projector P,



QR factorization

Let A be a $m \times n$ matrix $(m \geq n)$.

A QR factorization of A is a factorization of the form $A = QR$, where the columns of Q are orthonormal & R is upper Δ^r.

(reduced QR) $\quad \underset{m\times n}{\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix}} = \underset{m\times n}{\begin{bmatrix} q_1 & \cdots & q_n \\ & Q & \end{bmatrix}} \underset{R.\ n\times n}{\begin{bmatrix} * & & * \\ & \ddots & \\ 0 & & * \end{bmatrix}}$ $\qquad$ range A = range Q.

(full QR) $\quad \underset{m\times n}{\begin{bmatrix} & & \\ & A & \\ & & \end{bmatrix}} = \underset{m\times m}{\begin{bmatrix} q_1 & \cdots & q_n & q_{n+1} & \cdots & q_m \end{bmatrix}} \underset{m\times n}{\begin{bmatrix} * & & * \\ & \ddots & \\ 0 & & * \\ \hline & 0 & \end{bmatrix}}$

$\qquad\qquad\qquad Q \qquad\qquad\qquad\qquad R$

$$\text{range}(\{q_{n+1}, \cdots, q_m\}) = \text{range}(A^{\perp}) = \text{null}(A^*).$$

We will study 2 methods for computing the QR factorization.

(i) using Gram-Schmidt orthogonalization

(ii) Householder's matrices ; Givens matrices.

[I] Using Gram-Schmidt orthonormalization.

$A = QR$: $\quad \underset{m\times n}{\begin{bmatrix} a_1 & \cdots & a_n \end{bmatrix}} = \underset{Q}{\begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix}} \underset{R}{\begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1n} \\ 0 & r_{22} & & \vdots \\ \vdots & & \ddots & \\ 0 & & & r_{nn} \end{bmatrix}}$ $\qquad \bigstar$

G.S:
1) $\forall 1 \leq i \leq n$, $\langle a_1, \cdots, a_i \rangle = \langle q_1, \cdots, q_i \rangle$
2) $q_i^* q_j = \delta_{ij}$

$q_1 = \dfrac{a_1}{\|a_1\|}$, $\quad a_1 = q_1 \cdot \|a_1\|$, $\quad$ let $r_{11} = \|a_1\|$.

by G-S, $q_2 = \dfrac{a_2 - \langle q_1, a_2 \rangle q_1}{\| \cdots \|}$; from $\bigstar$: $a_2 = q_1 r_{12} + q_2 r_{22}$

i.e. $q_2 = \dfrac{a_2 - q_1 r_{12}}{r_{22}}$

$\therefore r_{12} = \langle q_1, a_2 \rangle$, $r_{22} = \| a_2 - \langle q_1, a_2 \rangle q_1 \|$

Continuing this way:

By ☆ $a_n = q_1 r_{1n} + q_2 r_{2n} + \cdots + \underline{q_n} r_{nn}$

$$\therefore q_n = \frac{a_n - q_1 r_{1n} - q_2 r_{2n} - \cdots - q_{n-1} r_{n-1,n}}{r_{nn}}$$

By G-S: $$q_n = \frac{a_n - q_1 \langle q_1, a_n \rangle - q_2 \langle q_2, a_n \rangle - \cdots - q_{n-1} \langle q_{n-1}, a_n \rangle}{\| \cdots \|}$$

$\therefore r_{in} = \langle q_i, a_n \rangle$ (for $i < n$) & $|r_{nn}| = \left\| a_n - \sum_{i=1}^{n-1} q_i \langle q_i, a_n \rangle \right\|$

In general, $r_{ij} = \langle q_i, a_j \rangle$ for $i < j$

$$\& \ |r_{jj}| = \left\| a_j - \sum_{i=1}^{j-1} q_i \langle q_i, a_j \rangle \right\|.$$

<u>Theorem</u>: Every $A \in \mathbb{C}^{m \times n}$ $(m \geq n)$ has a full QR factorization (hence also a reduced QR factorization)

<u>Pf</u>: If A has full rank, then G-S orthogonalization gives construction of QR factorization.

If A is not full rank, then for some j,

$$v_j = a_j - (q_1^* a_1) q_1 - \cdots - (q_{j-1}^* a_j) q_{j-1}$$

is zero. (In general, $q_i = \frac{v_i}{\|v_i\|}$ $1 \leq i \leq j-1$).

In this case, you can choose q_j arbitrarily such that

$\langle a_1, a_2, a_3 \rangle = \langle a_1, a_3 \rangle \subseteq \langle q_1, q_2, q_3 \rangle$

$$q_j \perp \langle q_1, \cdots, q_{j-1} \rangle \quad \left(v_{j+1} = a_{j+1} - \sum_{i=1}^{j} q_i^* a_{j+1} \right)$$

and continue the G-S process.

Uniqueness: If $A = QR$ is a reduced QR factorization, then multiplying ith column of Q by z & i-th row

$A = \overset{\downarrow}{Q} R \leftarrow \bar{z}^{1}$

of R by some scalar z s.t. $|z|=1$, this gives another QR factorization of A.

In particular, the sign of r_{jj} can be +ve or -ve.

If A is full rank & $r_{jj} > 0 \ \forall j$, then the QR factorization is unique.

rank $A = n-1$

$$\left[\begin{array}{c|c|c|c} a_1 & a_2 & \cdots & a_n \end{array}\right] = \begin{array}{c} \begin{array}{ccc} q_1 & q_2 & \overset{\downarrow z}{q_3} \end{array} \\ \left[\begin{array}{c|c|c|c} \frac{a_1}{\|a_1\|} & \frac{v_2}{\|v_2\|} & \frac{v_3}{\|v_3\|} & \cdots \end{array}\right] \end{array} \begin{bmatrix} \pm\|a_1\| & q_1^* a_2 & q_1^* a_3 & \\ & \pm\|v_2\| & q_2^* a_3 & \cdots \\ & & \pm\|v_3\| & \end{bmatrix} \leftarrow \bar{z}^{1}$$

$$v_2 = a_2 - \underbrace{\langle q_1, a_2\rangle}\, q_1 \quad , \quad q_2 = v_2/\|v_2\|$$

$$\rightarrow v_3 = a_3 - (q_1^* a_3) q_1 - (q_2^* a_3) q_2 \quad , \quad q_3 = v_3/\|v_3\|$$

$$\vdots$$

$$\rightarrow v_j = a_j - (\quad) q_1 - \cdots (\quad) q_{j-1} \quad , \quad q_j \perp \langle q_1, \ldots, q_{j-1}\rangle$$

$$q_{j+1} \perp \langle q_1, \ldots, q_j\rangle$$

$$\vdots$$

Algorithms

Classical G-S:

for $j = 1$ to n
  $v_j = a_j$
  for $i = 1$ to $j-1$
    $r_{ij} = q_i^* a_j$
    $v_j = v_j - r_{ij} q_i$
  end
  $r_{jj} = \|v_j\|_2$
  $q_j = v_j / r_{jj}$

Dry run:

$j=1$: $v_1 = a_1$, $r_{11} = \|a_1\|$, $q_1 = a_1/\|a_1\|$.

$j=2$: $i=1$: $r_{12} = q_1^* a_2$
$v_2 = v_2 - r_{12} q_1$
$r_{22} = \|v_2\|_2$
$q_2 = v_2 / r_{22}$.

$j=3$: $i=1$: r_{13}, $v_3 = v_3 - r_{13} q_1$
$i=2$: r_{23}, $v_3 = v_3 - r_{23} q_2$

Mathematically, the algorithm works as follows.

let $Q_{j-1} = \left[q_1 \middle| \cdots \middle| q_{j-1} \right]_{m\times(j-1)}$ $\quad \forall\, 2 \le j \le m$

Define $P_j = I - Q_{j-1} Q_{j-1}^*$, $P_1 = I$.

then P_j is the $m \times m$ matrix of rank $m-(j-1)$ which projects $\mathbb{C}^m$ orthogonally onto $\langle q_1, \dots, q_{j-1} \rangle^\perp$

$P_1 = I$, $P_2 = I - q_1 q_1^*$, $P_3 = I - Q_2 Q_2^*$, ...

Claim: $P_j a_j = a_j - r_{1j} q_1 - \cdots - r_{j-1\, j}\, q_{j-1}$.

$$P_j = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix}_{m \times m} - \begin{bmatrix} q_1 & \cdots & q_{j-1} \end{bmatrix} \begin{bmatrix} \cdots q_1^* \cdots \\ \vdots \\ q_{j-1}^* \end{bmatrix}$$

This computation is wrong. Find the correct computation at the last page.

Note that $$\left(Q_{j-1} Q_{j-1}^*\right)_{ij} = \begin{cases} 0 & \text{if } i \neq j \\ ii\text{-th component of } \left(\sum_1^{j-1} q_k q_k^*\right) & \text{if } i = j \end{cases}$$

$$P_j = I - Q_{j-1} Q_{j-1}^* = \begin{bmatrix} (1 - q_1 q_1^* - \cdots - q_{j-1} q_{j-1}^*)_1 & & \\ & \ddots & \\ & & (1 - q_1 q_1^* \cdots - q_{j-1} q_{j-1}^*)_i \\ & & & \ddots \end{bmatrix}_{m \times m}$$

However, the conclusion is true.

$$P_j a_j = a_j - q_1 (q_1^* a_j) - \cdots - q_{j-1} (q_{j-1}^* a_j)$$

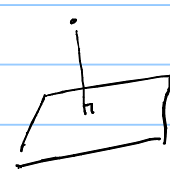
So, find r_j in the G-S process is equivalent to applying the projection P_j to a_j

$v_1 = P_1 a_1 \quad \to q_1$.

$v_2 = P_2 a_2 \quad \langle q_1 \rangle^\perp$

$v_3 = P_3 a_3 \quad \langle q_1, q_2 \rangle^\perp$

$\vdots$

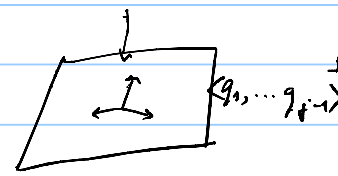


$v_j = P_j a_j \quad a_j \to \langle q_1, \cdots q_{j-1} \rangle^\perp$

The CGS can be viewed as computing at each step a single orthogonal projection of rank $m-(j-1)$.

The modified G-S algorithm computes the same result using a sequence of $j-1$ projections of rank $m-1$.

$$P_j := P_{\perp q_{j-1}} \cdots P_{\perp q_3} P_{\perp q_2} P_{\perp q_1}$$



$$P_j a_j = P_{\perp q_{j-1}} \cdots \left(P_{\perp q_2}\left(P_{\perp q_1} a_j\right)\right)$$

$$P_j = (I - q_{j-1}q_{j-1}^*) \cdots (I - q_2 q_2^*)(I - q_1 q_1^*)$$

$m \times m$.

rank. $m-(j-1)$

Check that this defn. of P_j gives the same as $P_j = I - Q_{j-1}Q_{j-1}^*$

Modified G-S: (layer-by-layer) or iteratively.

for $i = 1$ to n
  $v_i = a_i$.
for $i = 1$ to n
  $r_{ii} = \|v_i\|$
  $q_i = v_i / r_{ii}$
  for $j = i+1$ to n
    $r_{ij} = q_i^* v_j$
    $v_j = v_j - r_{ij} q_i$.

Dry run:

$r_{ii} = \|v_i\|$, $q_i = v_i / r_{ii}$

$i=1$: $r_{11} = \|v_1\|$, $q_1 = v_1 / r_{11}$,

r_{12} , $v_2^{(1)} = v_2 - r_{12} q_1$

r_{13} , $v_3^{(1)} = v_3 - r_{13} q_1$

$\vdots$

r_{1n} $v_n^{(1)} = v_n - r_{1n} q_n$

$i=2$: $r_{22} = \|v_2^{(1)}\|$, $q_2 = v_2^{(1)} / r_{22}$

r_{23} , $v_3^{(2)}$, & so on.

$$\begin{bmatrix} \quad & \quad \end{bmatrix} = \begin{bmatrix} q_1 & \cdots & q_n \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & r_{13} & \cdots & r_{1n} \\ & r_{22} & r_{23} & \cdots & r_{2n} \\ & & r_{33} & \cdots & r_{3n} \\ & & & \ddots & \vdots \\ & & & & r_{nn} \end{bmatrix}$$

Verification of $P_j(v) = v - \sum \langle v, q_i \rangle q_i$

First, consider $Q_1 = \begin{bmatrix} q_1 \end{bmatrix} = \begin{bmatrix} q_{11} \\ q_{21} \\ \vdots \\ q_{m1} \end{bmatrix}$

(just 1 column)

Then $$Q_1 Q_1^* = q_1 q_1^* = \begin{array}{c} \\ \\ i- \\ \\ \end{array}\begin{pmatrix} q_{11} \\ \vdots \\ q_{i1} \\ q_{m1} \end{pmatrix} \begin{array}{c} \qquad\qquad\quad j \\ \begin{pmatrix} \bar{q}_{11} & \bar{q}_{21} & \bar{q}_{j1} \cdots & \bar{q}_{m1} \end{pmatrix} \end{array} = \begin{pmatrix} q_{11}\bar{q}_{11} & q_{11}\bar{q}_{21} \cdots & q_{11}\bar{q}_{m1} \\ q_{21}\bar{q}_{11} & \cdots & q_{21}\bar{q}_{m1} \\ \vdots & & \\ q_{m1}\bar{q}_{11} & \cdots & q_{m1}\bar{q}_{m1} \end{pmatrix}$$

= $m \times m$ matrix with i-jth element being $q_{i1}\bar{q}_{j1}$.

So $$P_2 = I - Q_1 Q_1^* = \begin{bmatrix} 1 - q_{11}\bar{q}_{11} & -q_{11}\bar{q}_{21} & \cdots & -q_{11}\bar{q}_{m1} \\ -q_{21}\bar{q}_{11} & 1 - q_{21}\bar{q}_{21} & \cdots & -q_{21}\bar{q}_{m1} \\ \vdots & & & \\ -q_{m1}\bar{q}_{11} & \cdots & \cdots & 1 - q_{m1}\bar{q}_{m1} \end{bmatrix}$$

For a vector $v_i = \begin{pmatrix} v_{1i} \\ \vdots \\ v_{mi} \end{pmatrix}$,

$$P_2 v_i = \begin{pmatrix} (1 - q_{11}\bar{q}_{11}) v_{1i} - q_{11}\bar{q}_{21} v_{2i} - \cdots - q_{11}\bar{q}_{m1} v_{mi} \\ \vdots \\ -q_{m1}\bar{q}_{11} v_{1i} - q_{m1}\bar{q}_{21} v_{2i} - \cdots - (1 - q_{m1}\bar{q}_{m1}) v_{mi} \end{pmatrix}$$

$$= \begin{pmatrix} v_{1i} - \sum_{j=1}^{m} q_{11}\bar{q}_{j1} v_{ji} \\ \vdots \\ v_{mi} - \sum_{j=1}^{m} q_{m1}\bar{q}_{j1} v_{j1} \end{pmatrix}$$

$$= \begin{pmatrix} v_{1i} \\ \vdots \\ v_{mi} \end{pmatrix} - \begin{pmatrix} \sum_{j=1}^{m} q_{11}\bar{q}_{j1} v_{j1} \\ \vdots \\ \sum_{j=1}^{m} q_{m1}\bar{q}_{j1} v_{j1} \end{pmatrix} = v - (q_1 q_1^*) v.$$

$\underbrace{\begin{pmatrix} v_{1i} \\ \vdots \\ v_{mi} \end{pmatrix}}_{v}$

$\hookrightarrow = q_1 q_1^* v.$

$$Q_1 Q_1^* v = \begin{pmatrix} q_{11}\bar{q}_{11} & q_{11}\bar{q}_{21} & \cdots & q_{11}\bar{q}_{m1} \\ q_{21}\bar{q}_{11} & \cdots & & q_{21}\bar{q}_{m1} \\ \vdots & & & \\ q_{m1}\bar{q}_{11} & & \cdots & q_{m1}\bar{q}_{m1} \end{pmatrix} \begin{pmatrix} v_{1i} \\ \vdots \\ v_{mi} \end{pmatrix} = \begin{pmatrix} q_{11}\bar{q}_{11} v_{1i} + \cdots + q_{11}\bar{q}_{m1} v_{mi} \\ \vdots \\ q_{m1}\bar{q}_{11} v_{1i} + \cdots + q_{m1}\bar{q}_{m1} v_{mi} \end{pmatrix}$$

$$= \begin{pmatrix} \sum_{j=1}^{m} q_{11}\bar{q}_{j1} v_{j1} \\ \vdots \\ \sum_{j=1}^{m} q_{m1}\bar{q}_{j1} v_{j1} \end{pmatrix}$$

same as written above.

Note that $(q_1 q_1^*) v = q_1 (q_1^* v)$ by associativity

So you may compute $q_1^* v$ first & then compute $q_1 (q_1^* v)$ to get the same result.

This is what lets us view the G-S computation in 2 ways.

$$v - q_1 q_1^* v$$

Usual (inner product view)

Projection matrix view.

$v - q_1(q_1^* v)$

i.e. $v - \langle v, q_1 \rangle q_1$

these are equal.

$P_2(v) = (I - Q_1 Q_1^*) v$

$= v - (q_1 q_1^*) v$.

Now consider $Q_{j-1} = \left[q_1 \middle| \cdots \middle| q_{j-1} \right]$

Then $$Q_{j-1} Q_{j-1}^* = \begin{bmatrix} q_{11} & \cdots & q_{1j-1} \\ q_{21} & & \vdots \\ \vdots & & \vdots \\ q_{m1} & & q_{mj-1} \end{bmatrix} \begin{bmatrix} \bar{q}_{11} & \cdots & \bar{q}_{m1} \\ \vdots & & \\ \bar{q}_{1j-1} & \cdots & \bar{q}_{mj-1} \end{bmatrix}$$

$$= \begin{bmatrix} \sum_{k=1}^{j-1} q_{1k}\bar{q}_{1k} & \sum_{k=1}^{j-1} q_{1k}\bar{q}_{2k} & \cdots & \sum_{k=1}^{j-1} q_{1k}\bar{q}_{mk} \\ \vdots & & & \\ \sum_{k=1}^{j-1} q_{mk}\bar{q}_{1k} & & \cdots & \sum_{1}^{j-1} q_{mk}\bar{q}_{mk} \end{bmatrix}$$

$$= \sum_{k=1}^{j-1} \begin{bmatrix} q_{1k}\bar{q}_{1k} & q_{1k}\bar{q}_{2k} & \cdots & q_{1k}\bar{q}_{mk} \\ \vdots & & & \\ q_{mk}\bar{q}_{1k} & q_{mk}\bar{q}_{2k} & \cdots & q_{mk}\bar{q}_{mk} \end{bmatrix}$$

$$= q_1 q_1^* + q_2 q_2^* + \cdots + q_{j-1} q_{j-1}^* \quad \left(\text{Each term is a } m\times m \text{ matrix}\right)$$

So $P_j(v) = \left(I - Q_{j-1} Q_{j-1}^*\right) v$

$$= v - \left(q_1 q_1^* + q_2 q_2^* + \cdots + q_{j-1} q_{j-1}^*\right) v$$

$$= v - q_1 q_1^* v - q_2 q_2^* v - \cdots - q_m q_m^* v, \quad \text{as expected.}$$