QR & SVD

* QR - factorization

  - $A = QR$, Q has orthonormal columns that span col(A)

    R is upper triangular and invertible with positive diagonal entries

5/9 * Any $m \times n$ matrix with independent columns has a QR factorization $A = QR$

  - Q is a matrix with orthonormal columns
  - R is an invertible square matrix that is upper triangular and has positive diagonal entries

① We use Gram-Schmidt on the columns of A to find an orthonormal basis for col(A)

② Either take $R = Q^T A$ or
$$R = \begin{bmatrix} \|\vec{f}_1\| & \vec{c}_2 \cdot \vec{q}_1 & \vec{c}_3 \cdot \vec{q}_1 & \cdots & \vec{c}_n \cdot \vec{q}_1 \\ 0 & \|\vec{f}_2\| & \vec{c}_3 \cdot \vec{q}_2 & \cdots & \vec{c}_n \cdot \vec{q}_2 \\ \vdots & \vdots & \vdots & \ddots & \\ 0 & 0 & 0 & & \|\vec{f}_n\| \end{bmatrix}$$

- Ex:
$$A = \begin{bmatrix} 2 & 2 \\ 1 & -3 \\ -2 & -1 \end{bmatrix}$$
$\vec{c}_1$ $\vec{c}_2$

$\vec{f}_1 = \langle 2, 1, -2 \rangle$

$$\vec{f}_2 = \langle 2, -3, -1 \rangle - \frac{\langle 2,-3,-1\rangle \cdot \langle 2,1,-2\rangle}{\langle 2,1,-2\rangle \cdot \langle 2,1,-2\rangle} \langle 2,1,-2 \rangle$$
$$= \langle 2,-3,-1\rangle - \tfrac{3}{9}\langle 2,1,-2\rangle = \langle \tfrac{4}{3}, \tfrac{-10}{3}, \tfrac{-1}{3} \rangle$$

Check: $\langle 2,1,-2\rangle \cdot \langle \frac{4}{3}, \frac{-10}{3}, -\frac{1}{3} \rangle = \frac{8}{3} - \frac{10}{3} + \frac{2}{3} = 0$ ✓ (is orthogonal)

↳ $\left\{ \begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}, \begin{bmatrix} 4/3 \\ -10/3 \\ -1/3 \end{bmatrix} \right\}$ is an orthogonal basis for col(A)

↳ $\|\vec{f}_1\| = \sqrt{2^2 + 1^2 + (-2)^2} = \sqrt{9} = 3$ ⇝ $\vec{q}_1 = \frac{\vec{f}_1}{\|\vec{f}_1\|} = \langle \frac{2}{3}, \frac{1}{3}, \frac{-2}{3} \rangle$

$\|\vec{f}_2\| = \sqrt{16/9 + 100/9 + 1/9} = \sqrt{\frac{117}{9}} = \sqrt{13}$ ⇝ $\vec{q}_2 = \frac{\vec{f}_2}{\|\vec{f}_2\|} = \langle \frac{4}{3\sqrt{13}}, \frac{-10}{3\sqrt{13}}, \frac{-1}{3\sqrt{13}} \rangle$

↳ $Q = \begin{bmatrix} 2/3 & 4/3\sqrt{13} \\ 1/3 & -10/3\sqrt{13} \\ -2/3 & -1/3\sqrt{13} \end{bmatrix}$ orthonormal, $\vec{q}_1$, $\vec{q}_2$

$$R = Q^T A = \begin{bmatrix} 2/3 & 1/3 & -2/3 \\ 4/3\sqrt{13} & -10/3\sqrt{13} & -1/3\sqrt{13} \end{bmatrix} \begin{bmatrix} 2 & 2 \\ 1 & -3 \\ -2 & -1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 0 & \sqrt{13} \end{bmatrix}$$

* $QQ^T = I$ because Q is orthonormal and dot products will give 0

- What are the coordinates of the columns of A with respect to the basis in Q?

  First column: $\begin{bmatrix} 2 \\ 1 \\ -2 \end{bmatrix}$ has coordinates $\begin{bmatrix} 3 \\ 0 \end{bmatrix}$ (basically $3\vec{q}_1 + 0\vec{q}_2 = \vec{c}_1$)

  Second column: $\begin{bmatrix} 2 \\ -3 \\ -1 \end{bmatrix}$ has coordinates $\begin{bmatrix} 1 \\ \sqrt{13} \end{bmatrix}$ (basically $1\vec{q}_1 + \sqrt{13}\,\vec{q}_2 = \vec{c}_2$)

- Why is QR factorization useful?

$A^T A \vec{z} = A^T \vec{b}$ If $A = QR$

$A^T = (QR)^T = R^T Q^T$

$R^T \underbrace{Q^T Q}_{\text{identity matrix } I} R \vec{z} = R^T Q^T \vec{b}$

$R^T R \vec{z} = R^T Q^T \vec{b}$

$(R^T)^{-1} R^T R \vec{z} = (R^T)^{-1} R^T Q^T \vec{b}$

$\rightarrow R\vec{z} = Q^T \vec{b}$ ← R is upper triangular so solve with back-substitution

- Find QR factorization of $R_1 Q_1^T$

$R_1 Q_1^T = Q_2 R_2$
$R_2 Q_2^T = Q_3 R_3$ etc

} R converges to a matrix with the eigenvalues of A on the diagonal (since finding roots for 5+ degree polynomials is very difficult)

Q: when does A have a set of n orthonormal eigenvectors?

A: when A is symmetric

- $P^{-1}AP = P^TAP$ is diagonalization
- A^TA and AA^T have real and non-negative eigenvalues
- A^TA and AA^T have the same set of positive eigenvalues

* The singular values of A are the square root of the eigenvalues of A^TA

$\sigma_i = \sqrt{\lambda_i}$ ⇝ $\sigma_1 \geq \sigma_2 \geq \sigma_3 \geq \dots \geq \sigma_r$ $\sigma_i = 0$ for $i > r$

$$\Sigma_A = \begin{bmatrix} \sigma_1 & 0 & 0 & \cdots & & \\ 0 & \sigma_2 & 0 & \cdots & & \\ 0 & 0 & \ddots & & & \\ & & & \sigma_r & & \\ \vdots & \vdots & & & 0 & \\ & & & & & 0 \end{bmatrix}$$ ← singular matrix of A

- A^TA is symmetric → It has an orthonormal set of n eigenvectors
- $Q = [\vec{q}_1 \; \vec{q}_2 \; \cdots \; \vec{q}_n]$ where $\vec{q}_i$ is an eigenvector corresponding to σ_i
- $\{A\vec{q}_1, A\vec{q}_2, \dots, A\vec{q}_r\}$ is an orthogonal basis for col(A)

↳ $\vec{p}_1 = \frac{A\vec{q}_1}{\|A\vec{q}_1\|}$ $\vec{p}_2 = \frac{A\vec{q}_2}{\|A\vec{q}_2\|}$ $\dots$ $\vec{p}_r = \frac{A\vec{q}_r}{\|A\vec{q}_r\|}$

$\{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_r\}$ is an orthonormal basis for col(A)

(Use Gram-Schmidt to complete the basis $\{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_r, \vec{p}_{r+1}, \dots, \vec{p}_m\}$ is an orthonormal basis for $\mathbb{R}^m$)

* $A = P\Sigma_A Q^T$ (singular value decomposition, SVD), where P and Q are orthogonal matrices

rank(A) = r (# of nonzero σ)

$\{\vec{p}_1, \vec{p}_2, \dots, \vec{p}_r\}$ is an orthonormal basis for col(A)
$\{\vec{q}_1, \vec{q}_2, \dots, \vec{q}_r\}$ is an orthonormal basis for row(A)
$\{\vec{q}_{r+1}, \vec{q}_{r+2}, \dots, \vec{q}_n\}$ is an orthonormal basis for null(A)
$\{\vec{p}_{r+1}, \vec{p}_{r+2}, \dots, \vec{p}_m\}$ is an orthonormal basis for null(A^T)

- Rank 1 approximation to A :

$$P \begin{bmatrix} \sigma_1 & \cdots & \\ 0 & \ddots & \\ 0 & & \ddots \end{bmatrix} Q^T$$

$$\sigma_1 \vec{P} \begin{bmatrix} \vec{q}_1 \\ 0 \\ 0 \end{bmatrix}$$

$$\begin{array}{cc} & \begin{bmatrix} q_a & q_b & q_c & \cdots \end{bmatrix} \\ \begin{bmatrix} p_a \\ p_b \\ p_c \\ \vdots \end{bmatrix} & \begin{bmatrix} \sigma_1 p_a q_a & \sigma_1 p_b q_b & \cdots \\ \vdots & & \end{bmatrix} \end{array}$$

* Pseudo-inverse of A is A^+

↳ $A A^+ A = A$

$A^+ A A^+ = A^+$

$A^+ A$ and $A^+ A$ are symmetric

$A^+ = Q \Sigma'_A P^T$, $\Sigma'_A = \begin{bmatrix} 1/\sigma_1 & 0 & \cdots \\ 0 & 1/\sigma_2 & \\ \vdots & & \ddots \end{bmatrix}$