Finally solving $Rx = Q^T b$ finds the minimum of $\|Ax-b\|$

11. The QR factorization of a matrix A is given by

$A = QR$

$$Q = \begin{bmatrix} \frac{1}{3} & \frac{2}{3\sqrt{5}} \\ -\frac{2}{3} & \frac{\sqrt{5}}{3} \\ \frac{2}{3} & \frac{4}{3\sqrt{5}} \end{bmatrix} \quad \text{and} \quad R = \begin{bmatrix} 3 & \frac{1}{3} \\ 0 & \frac{4\sqrt{5}}{3} \end{bmatrix}.$$

Explain how to use this factorization to minimize $\|Ax - b\|$ and then find the minimizing value of x corresponding to $b = (1, 0, 1)$.

just solve $\boxed{Rx = Q^T b}$ to find the minimum

Now do it

$$\begin{bmatrix} \frac{1}{3} & \frac{2}{3\sqrt{5}} \\ -\frac{2}{3} & \frac{\sqrt{5}}{3} \\ \frac{2}{3} & \frac{4}{3\sqrt{5}} \end{bmatrix}^T = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3\sqrt{5} & \sqrt{5}/3 & 4/3\sqrt{5} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

Now solve

$$\begin{bmatrix} 3 & \frac{1}{3} \\ 0 & \frac{4\sqrt{5}}{3} \end{bmatrix} . x = \begin{bmatrix} 1/3 & -2/3 & 2/3 \\ 2/3\sqrt{5} & \sqrt{5}/3 & 4/3\sqrt{5} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}$$

by back substitution ☑

Don't do what follows: Find $A = \begin{bmatrix} \frac{1}{3} & \frac{2}{3\sqrt{5}} \\ -\frac{2}{3} & \frac{\sqrt{5}}{3} \\ \frac{2}{3} & \frac{4}{3\sqrt{5}} \end{bmatrix} \begin{bmatrix} 3 & \frac{1}{3} \\ 0 & \frac{4\sqrt{5}}{3} \end{bmatrix}.$

Then solve $A^T A x = A^T b$ the normal equations.

7.1 Diagonalization of Symmetric Matrices

Revisit the eigenvalue-eigenvector problem when $A^T = A$.

Claim if λ_1 and λ_2 are different eigenvalues and x_1 and x_2 the corresponding eigenvectors

Then $x_1 \cdot x_2 = 0$. Why?

$$Ax_1 = \lambda_1 x_1 \qquad\qquad Ax_2 = \lambda_2 x_2$$

$$\boxed{Ax_1 \cdot x_2} = \lambda_1 x_1 \cdot x_2 \qquad\qquad Ax_2 \cdot x_1 = \lambda_2 x_2 \cdot x_1$$

$$x_2 \cdot A^T x_1 = \lambda_2 x_1 \cdot x_2$$

same

symmetry $A^T = A$

$$\boxed{x_2 \cdot A x_1} = \lambda_2 x_1 \cdot x_2$$

Thus

$$\lambda_1 x_1 \cdot x_2 = \lambda_2 x_1 \cdot x_2$$

$$(\lambda_1 - \lambda_2)(x_1 \cdot x_2) = 0$$

Since $\lambda_1 \neq \lambda_2$ then $\lambda_1 - \lambda_2 \neq 0$ so $x_1 \cdot x_2 = 0$

This means one can make an orthogonal basis of eigenvector... one problem, what if x_i and λ_i are complex?

This can't happen

Claim if $A = A^T$ and $Ax = \lambda x$ then $x \in \mathbb{R}^n$ and $\lambda \in \mathbb{R}$

x is an eigenvector so $x \neq 0$.

Trying to show $\lambda \in \mathbb{R}$ equivalently $\bar{\lambda} = \lambda$

For contradiction suppose $\bar{\lambda} \neq \lambda$

$$Ax = \lambda x$$

$$\overline{Ax} = \overline{\lambda x}$$

$$\bar{A}\bar{x} = \bar{\lambda}\bar{x}$$ since $A \in \mathbb{R}^{n \times n}$ then $\bar{A} = A$

$$A\bar{x} = \bar{\lambda}\bar{x}$$ thus $\bar{\lambda}$ is an eigenvalue of A with eigenvector $\bar{x}$.

Since λ and $\bar{\lambda}$ are different eigenvalues

and x and $\bar{x}$ the corresponding eigenvectors

Then $x \cdot \bar{x} = 0$.

$$x \cdot \bar{x} = \sum_{i=1}^{n} x_i \bar{x_i} = \sum_{i=1}^{n} |x_i|^2 = 0$$

This means $x = 0$. But $x \neq 0$ since it's an eigenvector.

The conclusion λ and $\bar{\lambda}$ are not different eigenvalues

Thus $\lambda = \bar{\lambda}$ so $\lambda \in \mathbb{R}$.

Note also we can take x to be real.

$$Ax = \lambda x$$

$$A\bar{x} = \lambda\bar{x}$$

$$A(x + \bar{x}) = \lambda(x + \bar{x})$$

Even if x were not real then $v = x + \bar{x}$ would be real.
the only way v could be zero is if x were pure imaginary.

In that case $v = \frac{x}{i}$ would be real and also an eigenvector for λ.

The Spectral Theorem for Symmetric Matrices.

If $A \in \mathbb{R}^{n \times n}$ and $A^T = A$ then there is an orthonormal basis of eigenvectors with real eigenvalues

Thus $Ax_i = \lambda_i x_i$ where $x_i \in \mathbb{R}^n$, $x_i \neq 0$ and $\lambda_i \in \mathbb{R}$ for $i = 1, 2, \dots, n$.

and $$x_i \cdot x_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

The main thing to worry about is if there are multiple linearly independ eigenvectors for the same eigenvalue. For these eigenvectors use Gram-Schmidt to obtain an orthogonal (after normalization) orthonormal basis.

Suppose $x_1, x_2, \dots, x_p$ all had the same eigenvalue λ but they were linearly independent

$$\begin{bmatrix} \vdots & \vdots & & \vdots \\ x_1 & x_2 & \cdots & x_p \\ \vdots & \vdots & & \vdots \end{bmatrix} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ q_1 & q_2 & \cdots & q_p \\ \vdots & \vdots & & \vdots \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ & r_{22} & \cdots & r_{2p} \\ & 0 & \ddots & \vdots \\ & & & r_{pp} \end{bmatrix}$$

Note this equation says the x_i's are linear combinations of the q_i's.

$$\begin{bmatrix} \vdots & \vdots & & \vdots \\ x_1 & x_2 & \cdots & x_p \\ \vdots & \vdots & & \vdots \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ & r_{22} & \cdots & r_{2p} \\ & 0 & \ddots & \vdots \\ & & & r_{pp} \end{bmatrix}^{-1} = \begin{bmatrix} \vdots & \vdots & & \vdots \\ q_1 & q_2 & \cdots & q_p \\ \vdots & \vdots & & \vdots \end{bmatrix}$$

This says q_i's are linear combinations of the x_i's

$$A q_i = A(c_1 x_1 + c_2 x_2 + \cdots + c_p x_p)$$

$$= c_1 A x_1 + c_2 A x_2 + \cdots + c_p A x_p =$$

$$= c_1 \lambda x_1 + c_2 \lambda x_2 + \cdots + c_p \lambda x_p$$

$$= \lambda (c_1 x_1 + c_2 x_2 + \cdots + c_p x_p) = \lambda q_i$$

So q_i is also an eigenvector with eigenvalue λ.

What does the corresponding matrix factorization look like?

Since $A \in \mathbb{R}^{n \times n}$ and $A^T = A$. Then there are eigenvectors and eigenvalues so that

$$A x_i = \lambda_i x_i \text{ where } \underline{x_i \in \mathbb{R}^n},\ x_i \neq 0 \text{ and } \lambda_i \in \mathbb{R} \text{ for } i = 1, 2, \ldots, n.$$

and $$x_i \cdot x_j = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$$

$$Q = \begin{bmatrix} x_1 & x_2 & \cdots & x_n \end{bmatrix} \in \mathbb{R}^{n \times n} \qquad Q^T Q = \begin{bmatrix} x_1 \cdot x_1 & x_1 \cdot x_2 & \cdots & x_1 \cdot x_n \\ \vdots & & & 0 \\ x_n \cdot x_1 & x_n \cdot x_2 & \cdots & x_n \cdot x_n \end{bmatrix} = I$$

Since $Q^T Q = I$ and Q is square then $Q^{-1} = Q^T$.
and Q is an orthogonal matrix

$$AQ = QD \quad \text{where } D = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_3 \end{bmatrix}$$

$$A = QDQ^{-1} = QDQ^T$$

Factorization $A = QDQ^T$ when $A = A^T$

orthogonal matrix — diagonal matrix — orthogonal matrix

Monday is singular value decomposition:

Idea what to do when A is not symmetric.
and how to use the spectral theorem.

$B = A^T A$ then $B^T = (A^T A)^T = A^T A^{TT} = A^T A = B$

$C = AA^T$ then $C^T = (AA^T)^T = A^{TT} A^T = AA^T = C$

It could happen $A \in \mathbb{R}^{m \times n}$

$B = \underset{n \times m}{A^T} \underset{m \times n}{A} \in \mathbb{R}^{n \times n}$ $C = \underset{m \times n}{A} \underset{n \times m}{A^T} \in \mathbb{R}^{m \times m}$

If $m \neq n$ then choose the matrix B or C which is smaller.