Lecture 14

Application of GS.: QR decomposition.

- QR decomposition is a technique for decomposing certain matrices in a nice way.
- Useful in many numerical applications

<u>Question</u>: Let A be an $m \times n$ matrix with linearly independent column vectors. Let Q be the matrix obtained by applying G.S. to the columns of A. How are A and Q related?

- Let $A = [\vec{v}_1 | \vec{v}_2 | \dots | \vec{v}_n]$ and let

$$Q = [\vec{q}_1 | \vec{q}_2 | \dots | \vec{q}_n]$$

where $\{\vec{q}_1, \dots \vec{q}_n\}$ is the ortho-normal set obtained by applying Gram Schmidt to $\{\vec{v}_1, \dots \vec{v}_n\}$.

- Recall that

$$\vec{v}_1 = \langle \vec{v}_1, \vec{q}_1 \rangle \vec{q}_1 + \langle \vec{v}_1, \vec{q}_2 \rangle \vec{q}_2 + \dots + \langle \vec{v}_1, \vec{q}_n \rangle \vec{q}_n.$$

$$\vec{v}_2 = \langle \vec{v}_2, \vec{q}_1 \rangle \vec{q}_1 + \langle \vec{v}_2, \vec{q}_2 \rangle \vec{q}_2 + \dots + \langle \vec{v}_2, \vec{q}_n \rangle \vec{q}_n$$

$$\vdots$$

$$\vec{v}_n = \langle \vec{v}_n, \vec{q}_1 \rangle \vec{q}_1 + \langle \vec{v}_n, \vec{q}_2 \rangle \vec{q}_2 + \dots + \langle \vec{v}_n, \vec{q}_n \rangle \vec{q}_n..$$

So

$$\underbrace{\left[\vec{v}_1 \,|\, \vec{v}_2 \,|\, \cdots \,|\, \vec{v}_n\right]}_{A} = \underbrace{\left[\vec{q}_1 \,|\, \vec{q}_2 \,|\, \cdots \,|\, q_n\right]}_{Q} \underbrace{\begin{bmatrix} \langle \vec{v}_1, \vec{q}_1 \rangle & \cdots\cdots & \langle \vec{v}_n, \vec{q}_1 \rangle \\ \langle \vec{v}_1, \vec{q}_2 \rangle & \cdots\cdots & \langle \vec{v}_n, \vec{q}_2 \rangle \\ \vdots & & \vdots \\ \langle \vec{v}_1, q_n \rangle & \cdots & \langle \vec{v}_n, \vec{q}_n \rangle \end{bmatrix}}_{R}.$$

- Note that R is an $n \times n$ matrix.
- However, by the Gram-Schmidt construction,

  $\langle v_i, q_j \rangle = 0$ for all $i \lneq j$ and so.

$$R = \begin{bmatrix} \langle \vec{v}_1, \vec{q}_1 \rangle & \langle \vec{v}_2, \vec{q}_1 \rangle & \cdots & \langle \vec{v}_n, q_1 \rangle \\ 0 & \langle \vec{v}_2, \vec{q}_2 \rangle & \cdots & \langle \vec{v}_2, q \rangle \\ \vdots & 0 & & \vdots \\ \vdots & \vdots & & \langle \vec{v}_n, q_n \rangle \\ 0 & 0 & \cdots & \end{bmatrix}.$$

an upper triangular square matrix.

- Note also, that for all i, $\langle \vec{v}_i, \vec{q}_i \rangle \neq 0$

  and so R is necessarily <u>invertible</u>.

- $A = QR$ is called the QR-decomposition of A.

Theorem: (QR Decomposition) If A is an $m \times n$ matrix with linearly independent columns then A can be factored as $A = QR$ where Q is an $m \times n$ matrix with orthonormal column vectors, and R is an invertible upper triangular matrix.

Example: Let $A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix}$.

$$\vec{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 1 \end{bmatrix}, \quad \vec{v}_3 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}.$$

Apply GS with normalization to set

$$\vec{q}_1 = \begin{bmatrix} 1/\sqrt{3} \\ 1/\sqrt{3} \\ 1/\sqrt{3} \end{bmatrix} \quad \vec{q}_2 = \begin{bmatrix} -2/\sqrt{6} \\ 1/\sqrt{6} \\ 1/\sqrt{6} \end{bmatrix} \quad \vec{q}_3 = \begin{bmatrix} 0 \\ -1/\sqrt{2} \\ 1/\sqrt{2} \end{bmatrix}.$$

Now,
$$R = \begin{bmatrix} \langle \vec{v}_1, \vec{q}_1 \rangle & \langle \vec{v}_2, \vec{q}_1 \rangle & \langle \vec{v}_3, \vec{q}_1 \rangle \\ 0 & \langle \vec{v}_2, \vec{q}_2 \rangle & \langle \vec{v}_3, \vec{q}_2 \rangle \\ 0 & 0 & \langle \vec{v}_3, \vec{q}_3 \rangle \end{bmatrix} = \begin{bmatrix} 3/\sqrt{3} & 2/\sqrt{3} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{6} \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix}.$$

So
$$A = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 1 & 1 & 1 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} & -2/\sqrt{6} & 0 \\ 1/\sqrt{3} & 1/\sqrt{6} & -1/\sqrt{2} \\ 1/\sqrt{3} & 1/\sqrt{6} & 1/\sqrt{2} \end{bmatrix} \begin{bmatrix} 3/\sqrt{3} & 2/\sqrt{3} & 1/\sqrt{3} \\ 0 & 2/\sqrt{6} & 1/\sqrt{6} \\ 0 & 0 & 1/\sqrt{2} \end{bmatrix}.$$

Least Squares approximation.

- Suppose we have a linear system

$$A\vec{x} = \vec{b}$$

which we know to be inconsistent.

- Since we cannot find an exact solution we want to find a solution which is as good as possible.

- This means that we want to make the error as small as possible, the error being

$$\|\vec{b} - A\vec{x}\|.$$

- Such a solution is called a "least squares" solution, since if we write

$$\vec{b} - A\vec{x} = \begin{bmatrix} e_1 \\ \vdots \\ e_n \end{bmatrix} \leftarrow \text{"error."}$$

then minimizing $\|\vec{b} - A\vec{x}\|$ (wrt the standard inner product) is equivalent to minimizing

$$e_1^2 + e_2^2 + \dots + e_n^2.$$

- **Note:** For every vector $\vec{x} \in \mathbb{R}^n$, $A\vec{x}$ is in the column space of A (i.e.

a linear combo of the columns of A, $col(A)$).

- Therefore, a least squares solution of $A\vec{x}=\vec{b}$ is a vector $A\hat{x} \in col(A)$ which is as close to $\vec{b}$ as possible.

- It turns out this is always the projection of $\vec{b}$ onto $col(A)$ ie.

$$A\hat{x} = proj_{col(A)}\vec{b}.$$

<u>Theorem</u> (Best Approximation)

Let W be a finite dimensional subspace of an inner product space V. For any $\vec{b} \in V$, $proj_W \vec{b}$ is the <u>best approximation of $\vec{b}$ from W</u> in the sense that

$$\|\vec{b} - proj_W \vec{b}\| < \|\vec{b} - \vec{w}\|$$

for every $\vec{w} \in W$ such that $\vec{w} \neq proj_W \vec{b}$.

<u>Pf:</u> Let $\vec{w} \in W$ be arbitrary. Then we can write

$$\vec{b} - \vec{w} = (\vec{b} - proj_W \vec{b}) + (\underbrace{proj_W \vec{b} - \vec{w}}_{\in W}).$$

- Now $proj_W \vec{b} \in W$ and $\vec{w} \in W$, so $proj_W \vec{b} - \vec{w} \in W$

- Also, $\vec{b} - proj_W \vec{b} \in W^{\perp}$, so $\vec{b} - proj_W \vec{b}$ and $proj_W \vec{b} - \vec{w}$ are orthogonal.

- By the generalized Pythagorean theorem,

$$\|\vec{b}-\vec{w}\|^2 = \|\vec{b}-\mathrm{proj}_W\vec{b}\|^2 + \|\mathrm{proj}_W\vec{b}-\vec{w}\|^2.$$

If $\vec{w} \neq \mathrm{proj}_W\vec{b}$, then $\|\mathrm{proj}_W\vec{b}-\vec{w}\|^2 > 0$,

so then

$$\|\vec{b}-\mathrm{proj}_W\vec{b}\|^2 < \|\vec{b}-\vec{w}\|^2,$$

and hence

$$\|\vec{b}-\mathrm{proj}_W\vec{b}\| < \|\vec{b}-\vec{w}\|$$

as required. $\square$.