July 19

## Gram-Schmidt process

Yesterday we defined
- Orthogonal and orthonormal bases
- $\text{proj}_W$ when we have such a basis

But does every subspace of $\mathbb{R}^n$ even have one?

The Gram-Schmidt process is an algorithm which, given a basis $\vec{v}_1, \dots, \vec{v}_p$ of a subspace $W$, produces an orthogonal basis $\vec{u}_1, \dots, \vec{u}_p$ step-by-step. It satisfies the following defining rules: For $1 \leq n \leq p$,

1) $\vec{u}_1, \dots, \vec{u}_n$ is an orthogonal set
2) $\text{Span}\{\vec{u}_1, \dots, \vec{u}_n\} = \text{Span}\{\vec{v}_1, \dots, \vec{v}_n\}$

These rules are easily implemented. Let $W_n = \text{Span}\{\vec{u}_1, \dots, \vec{u}_n\}$. Then let $\vec{u}_{n+1} = \vec{v}_{n+1} - \text{proj}_{W_n} \vec{v}_{n+1}$ (the component of $\vec{v}_{n+1}$ in $W_n^\perp$, so rule 1 is satisfied).

For rule 2, suppose we have shown $\text{Span}\{\vec{u}_1, \dots, \vec{u}_n\} = \text{Span}\{\vec{v}_1, \dots, \vec{v}_n\}$.

$$\begin{aligned}\text{Span}\{\vec{u}_1, \dots, \vec{u}_{n+1}\} &= \text{Span}\{\vec{u}_1, \dots, \vec{u}_n, \vec{v}_{n+1} - \text{proj}_{W_n}\vec{v}_{n+1}\} \\ &= \text{Span}\{\vec{u}_1, \dots, \vec{u}_n, \vec{v}_{n+1}\} \\ &= \text{Span}\{\vec{v}_1, \dots, \vec{v}_n, \vec{v}_{n+1}\}.\end{aligned}$$

Thus, by induction, rule 2 is satisfied.

ex $\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}1\\0\\0\end{bmatrix}, \begin{bmatrix}0\\1\\0\end{bmatrix}$ is basis of $\mathbb{R}^3$.

$\vec{u}_1 = \begin{bmatrix}1\\1\\1\end{bmatrix}$. $\quad \dfrac{\vec{u}_1 \cdot \begin{bmatrix}1\\0\\0\end{bmatrix}}{\vec{u}_1 \cdot \vec{u}_1} = \dfrac{1}{3} \quad \vec{u}_2 = \begin{bmatrix}1\\0\\0\end{bmatrix} - \frac{1}{3}\vec{u}_1 = \begin{bmatrix}2/3\\-1/3\\-1/3\end{bmatrix}$

$\dfrac{\vec{u}_1 \cdot \begin{bmatrix}0\\1\\0\end{bmatrix}}{\vec{u}_1 \cdot \vec{u}_1} = \dfrac{1}{3} \quad \dfrac{\vec{u}_2 \cdot \begin{bmatrix}0\\1\\0\end{bmatrix}}{\vec{u}_2 \cdot \vec{u}_2} = \dfrac{-1/3}{6/9} = -\dfrac{1}{2} \quad \vec{u}_3 = \begin{bmatrix}0\\1\\0\end{bmatrix} - \frac{1}{3}\vec{u}_1 + \frac{1}{2}\vec{u}_2 = \begin{bmatrix}0\\1/2\\-1/2\end{bmatrix}$

so $\begin{bmatrix}1\\1\\1\end{bmatrix}, \begin{bmatrix}2/3\\-1/3\\-1/3\end{bmatrix}, \begin{bmatrix}0\\1/2\\-1/2\end{bmatrix}$ is an orthogonal basis.

Can make orthonormal by normalizing.

ex Orthonormal basis from $\begin{bmatrix}3\\6\\0\end{bmatrix}, \begin{bmatrix}1\\2\\2\end{bmatrix}$.

$\vec{u}_1 = \begin{bmatrix}3\\6\\0\end{bmatrix}$ $\qquad \dfrac{\vec{u}_1 \cdot \begin{bmatrix}1\\2\\2\end{bmatrix}}{\vec{u}_1 \cdot \vec{u}_1} = \dfrac{15}{45} = \dfrac{1}{3}$ $\qquad \begin{bmatrix}1\\2\\2\end{bmatrix} - \frac{1}{3}\vec{u}_1 = \begin{bmatrix}0\\0\\2\end{bmatrix}$

$\frac{1}{\sqrt{45}}\begin{bmatrix}3\\6\\0\end{bmatrix}, \frac{1}{2}\begin{bmatrix}0\\0\\2\end{bmatrix}$ is orthonormal basis.

ex $\vec{v}_1, \vec{v}_2, \vec{v}_3$ span $\mathbb{R}^3$. Find an orthogonal basis.

No need for Gram-Schmidt: $\vec{e}_1, \vec{e}_2, \vec{e}_3$

## QR factorization

Suppose $A$ is $m \times n$ with linearly independent columns.
A QR factorization of $A$ is

$$A = QR$$

with $Q$ $m \times n$ having orthonormal columns and
$R$ $n \times n$ upper triangular (invertible)
with positive entries on the diagonal.

$\text{Col } A = \text{Col } QR = \text{Col } Q$ (since $R$ is onto)
That is, the columns of $Q$ are an orthonormal basis of $\text{Col } A$.

In practice, QR factorizations are computed with Givens rotations, which for some reason are not in the curriculum. Basically, you "rotate" rows of $A$ to obtain an upper triangular matrix. It is more numerically stable to do this than Gaussian elimination (which is the LU factorization).

Why is QR factorization useful? Suppose you want an approximate solution to $A\vec{x}=\vec{b}$, that is, minimizing $\|A\hat{x}-\vec{b}\|$. $\quad \text{proj}_{\text{Col}\,A}\,\vec{b} = QQ^T\vec{b}$, so

$$\|A\hat{x}-\vec{b}\| \text{ minimized when } \|A\hat{x}-QQ^T\vec{b}\| \text{ is.}$$
$$= \|Q(R\hat{x}-Q^T\vec{b})\| = \|R\hat{x}-Q^T\vec{b}\|.$$

But $R\hat{x}=Q^T\vec{b}$ is solvable as $\hat{x}=R^{-1}Q^T\vec{b}$.

(Note: in many texts, $Q$ is $m\times m$ orthogonal and $R$ is $m\times n$ upper triangular. $\|A\hat{x}-\vec{b}\| = \|Q^T(A\hat{x}-\vec{b})\| = \|R\hat{x}-Q^T\vec{b}\|$.
Since $R$ is upper triangular,

$$[R \mid Q^T\vec{b}] = \left[\begin{array}{ccc|c} & \diagdown & & \\ 0 & & & (Q^T\vec{b})_{1\cdots n} \\ \hline 0 & \cdots & 0 & \\ \vdots & & \vdots & (Q^T\vec{b})_{n+1\cdots m} \\ 0 & \cdots & 0 & \end{array}\right]$$

The upper part has a unique solution $\hat{x}$, then $\|R\hat{x}-Q^T\vec{b}\| = \|(Q^T\vec{b})_{n+1\cdots m}\|$ is the "residual".

Gram-Schmidt gives one way to obtain QR factorizations.

ex
$$\begin{bmatrix} 1 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} = \begin{bmatrix} 1/\sqrt{3} & 2/\sqrt{6} & 0 \\ 1/\sqrt{3} & -1/\sqrt{6} & 1/\sqrt{2} \\ 1/\sqrt{3} & -1/\sqrt{6} & -1/\sqrt{2} \end{bmatrix} R$$

(proj onto Col A)

Since $QQ^T\vec{v} = \vec{v}$ for $\vec{v} \in \text{Col } A$,

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

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

We can also obtain R straight from Gram-Schmidt.

ex $\begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 \\ 2 \\ 2 \end{bmatrix}$ had $\vec{u_1} = \begin{bmatrix} 3 \\ 6 \\ 0 \end{bmatrix}$ and $\vec{u_2} = \begin{bmatrix} 0 \\ 0 \\ 2 \end{bmatrix}$

with $\vec{u_1} = \vec{v_1}$ and $\vec{u_2} = \vec{v_2} - \frac{1}{3}\vec{u_1}$.

Since $\frac{1}{\sqrt{45}}\vec{u_1}, \frac{1}{2}\vec{u_2}$ is orthonormal,

$$\frac{1}{\sqrt{45}}\vec{u_1} = \frac{1}{\sqrt{45}}\vec{v_1} \Rightarrow \vec{v_1} = \sqrt{45}\left(\frac{1}{\sqrt{45}}\vec{u_1}\right)$$

$$\vec{v_2} = \frac{1}{3}\vec{u_1} + \vec{u_2} \Rightarrow \vec{v_2} = \frac{\sqrt{45}}{3}\left(\frac{1}{\sqrt{45}}\vec{u_1}\right) + 2\left(\frac{1}{2}\vec{u_2}\right)$$

so $$\begin{bmatrix} 3 & 1 \\ 6 & 2 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 3/\sqrt{45} & 0 \\ 6/\sqrt{45} & 0 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} \sqrt{45} & \sqrt{45}/3 \\ 0 & 2 \end{bmatrix}$$

is a QR factorization.

QR factorization is actually unique. The proof relies on positive diagonal entries for R.

# Least-Squares

Given an inconsistent system $A\vec{x} = \vec{b}$, we may be interested in $\hat{x}$ such that $\|A\hat{x} - \vec{b}\|$ is minimized.

Since $A\hat{x} \in \text{Col } A$, by "best approx. thm.",

$A\hat{x} = \text{proj}_{\text{Col } A} \vec{b}$ is best solution.

Alternatively, want $A\hat{x} - \vec{b}$ orthogonal to Col A.

So, ortho. to each column.

$$A^T(A\hat{x} - \vec{b}) = \vec{0} \Rightarrow A^TA\hat{x} = A^T\vec{b}.$$