Gram-Schmidt Orthogonalization Process

Mon. Apr. 27 | Math 304-6 Feingold

Th. Let $W \leq \mathbb{R}^n$ have basis $T = \{w_1, \dots, w_m\}$. Then an orthogonal basis $T' = \{w_1', \dots, w_m'\}$ for W can be found as follows:

$$w_1' = w_1 \quad , \quad w_2' = w_2 - \mathrm{Proj}_{w_1'}(w_2) = w_2 - \left(\frac{w_2 \cdot w_1'}{w_1' \cdot w_1'}\right) w_1' \quad ,$$

$$w_3' = w_3 - \mathrm{Proj}_{\langle w_1', w_2' \rangle}(w_3) =$$

$$w_3 - \left(\frac{w_3 \cdot w_1'}{w_1' \cdot w_1'}\right) w_1' - \left(\frac{w_3 \cdot w_2'}{w_2' \cdot w_2'}\right) w_2'$$

$\vdots$

$$w_i' = w_i - \mathrm{Proj}_{\langle w_1', \dots, w_{i-1}' \rangle}(w_i) = w_i - \sum_{j=1}^{i-1} \left(\frac{w_i \cdot w_j'}{w_j' \cdot w_j'}\right) w_j'$$

for $1 \leq i \leq m$. Also,

$$\langle w_1, \dots, w_i \rangle = \langle w_1', \dots, w_i' \rangle \quad \text{for } 1 \leq i \leq m.$$

Pf. We have defined $\mathrm{Proj}_{w_1'}(w_2)$ such that $(w_2 - \mathrm{Proj}_{w_1'}(w_2)) \perp w_1'$ so $w_2' \perp w_1'$ makes $\{w_1', w_2'\}$ an orthogonal basis for its span. But then $(w_3 - \mathrm{Proj}_{\langle w_1', w_2' \rangle}(w_3)) \perp \langle w_1', w_2' \rangle$ so $\{w_1', w_2', w_3'\}$ is an orthogonal basis for its span. The formula given for that projection was proved before based on the assumption that basis $\{w_1', w_2', w_3'\}$ is orthogonal. The formula for w_i' follows by the same argument (by induction) since $\{w_1', \dots, w_{i-1}'\}$ is an orthogonal basis of its span. That formula also shows that $w_i' \in \langle w_1', \dots, w_{i-1}', w_i \rangle = \langle w_1, \dots, w_{i-1}, w_i \rangle$ and $w_i \in \langle w_1', \dots, w_{i-1}', w_i' \rangle$ so

$\langle w_1, \dots, w_i \rangle = \langle w_1', \dots, w_i' \rangle$. $\square$

<u>Example</u>: Let $W \le \mathbb{R}^4$ where a basis of W is

$$T = \left\{ w_1 = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \; w_2 = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \; w_3 = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} \right\}.$$

Then G-S. process is:

$$w_1' = w_1, \quad w_2' = w_2 - \left(\frac{w_2 \cdot w_1'}{w_1' \cdot w_1'} \right) w_1' = \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} - \left(\frac{2}{2} \right) \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}$$

$$w_3' = w_3 - \left(\frac{w_3 \cdot w_1'}{w_1' \cdot w_1'} \right) w_1' - \left(\frac{w_3 \cdot w_2'}{w_2' \cdot w_2'} \right) w_2'$$

$$= \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} - \left(\frac{4}{2} \right) \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix} - \left(\frac{6}{2} \right) \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \\ 4 \end{bmatrix} - \begin{bmatrix} 2 \\ 0 \\ 2 \\ 0 \end{bmatrix} - \begin{bmatrix} 0 \\ 3 \\ 0 \\ 3 \end{bmatrix} = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}$$

so $T' = \left\{ w_1' = \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \; w_2' = \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \; w_3' = \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix} \right\}$ is an orthogonal basis of W. Check that $w_i' \cdot w_j' = 0$ for $i \ne j$.

<u>Th.</u> (Normalization Step of G.-S.)
After obtaining orthogonal basis T' from the G.-S. process, we can replace each vector $w_i' \in T'$ by a unit vector, $w_i'' = \frac{w_i'}{\|w_i'\|}$, to get an orthonormal basis $T'' = \{w_1'', \dots, w_m''\}$ of W.

<u>EX:</u> In the last example, $\|w_1'\| = \sqrt{2} = \|w_2'\|$ and $\|w_3'\| = \sqrt{4} = 2$ so an orthonormal basis of W is

$$T'' = \left\{ w_1'' = \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\1\\0\end{bmatrix}, w_2'' = \frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\\0\\1\end{bmatrix}, w_3'' = \frac{1}{2}\begin{bmatrix}-1\\-1\\1\\1\end{bmatrix} \right\}$$

<u>Exercise:</u> Use T' to find $x_1, x_2, x_3 \in \mathbb{R}$ s.t. $\mathrm{Proj}_W(v) = x_1 w_1' + x_2 w_2' + x_3 w_3'$ for any $v = \begin{bmatrix}a\\b\\c\\d\end{bmatrix} \in \mathbb{R}^4$.

<u>Example</u>: For $W = \langle T \rangle = \langle T' \rangle = \langle T'' \rangle$ as before, find $W^{\perp} = \{X \in \mathbb{R}^4 \mid X \perp W\}$ and use its basis vector to extend T' to an orthogonal basis of $\mathbb{R}^4$, extend T'' to an orthonormal basis of $\mathbb{R}^4$, and use that answer to give an orthogonal matrix $A \in \mathbb{R}^4_4$.

<u>Solution</u>: Find $W^{\perp} = \{X \in \mathbb{R}^4 \mid X \cdot w_i' = 0, 1 \le i \le 3\}$ by solving

$$\left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ -1 & -1 & 1 & 1 & 0 \end{array}\right] \to \left[\begin{array}{cccc|c} 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 2 & 2 & 0 \end{array}\right] \to \left[\begin{array}{cccc|c} 1 & 0 & 0 & -1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 0 \end{array}\right] \begin{array}{l} x_1 = r \\ x_2 = -r \\ x_3 = -r \\ x_4 = r \in \mathbb{R} \end{array}$$

$$W^{\perp} = \left\{ \begin{bmatrix} r \\ -r \\ -r \\ r \end{bmatrix} \in \mathbb{R}^4 \;\middle|\; r \in \mathbb{R} \right\} = \left\langle \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \right\rangle \text{ so}$$

$$S' = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \end{bmatrix}, \begin{bmatrix} 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} -1 \\ -1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix} \right\}$$ is an orthogonal basis of $\mathbb{R}^4$ extending T'.

Normalizing each vector in S' gives o.n. basis

$$S'' = \left\{ \frac{1}{\sqrt{2}}\begin{bmatrix}1\\0\\1\\0\end{bmatrix}, \frac{1}{\sqrt{2}}\begin{bmatrix}0\\1\\0\\1\end{bmatrix}, \frac{1}{2}\begin{bmatrix}-1\\-1\\1\\1\end{bmatrix}, \frac{1}{2}\begin{bmatrix}1\\-1\\-1\\1\end{bmatrix} \right\}$$ of $\mathbb{R}^4$, and

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

is an orthogonal matrix in $\mathbb{R}^4$ whose columns are the vectors in o.n. set S''.

Check that $AA^T = I_4 = A^TA$

Important results about orthog. matrices:

Th: For any $X, Y \in \mathbb{R}^n$ and for any $A \in \mathbb{R}^n_n$, we have $(AX)\cdot Y = X\cdot(A^T Y)$.

Pf. $(AX)\cdot Y = (AX)^T Y = (X^T A^T) Y = X^T (A^T Y) = X\cdot(A^T Y)$

Th: For any $X, Y \in \mathbb{R}^n$ if $A \in \mathbb{R}^n_n$ is orthogonal then $(AX)\cdot(AY) = X\cdot Y$.

Pf. Since A orthog. means $A^T = A^{-1}$, we have
$(AX)\cdot(AY) = X\cdot(A^T A Y) = X\cdot(A^{-1}AY) = X\cdot(I_n Y) = X\cdot Y$.

Cor: If $A^T = A^{-1}$ then for any $X, Y \in \mathbb{R}^n$ we have
$\|AX\| = \|X\|$ and $\theta_{X,Y} = \theta_{AX,AY}$ so $L_A : \mathbb{R}^n \to \mathbb{R}^n$ preserves lengths and angles.

Important result about symm. matrices: 8

Th: Let $A = A^T \in \mathbb{R}^n_n$ be symmetric and let $\lambda \neq \mu$ in $\mathbb{R}$ be e-values of A. Then the e-spaces A_λ and A_μ are perpendicular, $A_\lambda \perp A_\mu$.

Pf. Need to show that for any $X \in A_\lambda$, $Y \in A_\mu$ that $X \cdot Y = 0$. It is clear if $X = 0^n_1$ or $Y = 0^n_1$ so suppose $AX = \lambda X$ and $AY = \mu Y$ for $X, Y \in \mathbb{R}^n_1$ nonzero e-vectors. Then we have

$$\lambda(X \cdot Y) = (\lambda X) \cdot Y = (AX) \cdot Y = X \cdot (A^T Y) = X \cdot (AY) = X \cdot (\mu Y)$$
$$= \mu(X \cdot Y)$$

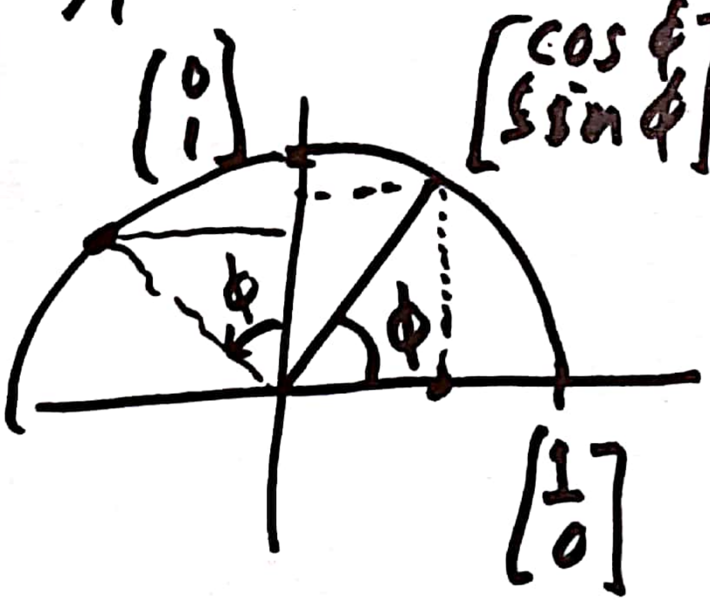
So $\lambda(X \cdot Y) = \mu(X \cdot Y)$ so $(\lambda - \mu)(X \cdot Y) = 0$. But $\lambda - \mu \neq 0$ so $\overline{X \cdot Y = 0}$ □

Th: Let $A = A^T \in \mathbb{R}^n_n$ and suppose $\lambda_1, \ldots, \lambda_r \in \mathbb{R}$ are the distinct e-values of A. Let T_i' be an orthonormal basis of e-space A_{λ_i}, $1 \le i \le r$, obtained by Gram-Schmidt process from any basis T_i of A_{λ_i}. Then $T' = T_1' \cup T_2' \cup \ldots \cup T_r'$ is an orthonormal basis of $\mathbb{R}^n$ and $P = {}_SP_{T'}$ is an orthogonal matrix (whose columns are the vectors in T') such that $P^{-1}AP = D$ is diagonal with blocks $\lambda_i I_{g_i}$ on the diagonal, $g_i = \dim(A_{\lambda_i}) = k_i$ (geom. = alg. mult.).

Since P is orthog. $P^{-1} = P^T$ so $D = P^TAP$ and we say A can be "orthogonally diag-ized."

Pf. In Advanced Lin. Alg. it is shown that all e-values of symm. $A \in \mathbb{R}^n_n$ are real, and that $g_i = k_i$ so A is diag-able. Since G.-S. gives orthonormal bases T_i' for each A_{λ_i}, and $A_{\lambda_i} \perp A_{\lambda_j}$ for $1 \le i \ne j \le r$ by last Theorem, we get that T' is an orthonormal set of n e-vectors in $\mathbb{R}^n$ so $P = {}_S P_{T'}$ is orthog., $P^{-1} = P^T$ and $D = P^T A P$ is diag. with the e-values λ_i on the diag. repeated $g_i = k_i$ times in blocks corresponding to the order of e-vectors in T'. $\square$

$L_A : \mathbb{R}^2 \to \mathbb{R}^2$ for $A = \begin{bmatrix} \cos\phi & -\sin\phi \\ \sin\phi & \cos\phi \end{bmatrix}$



$L_A \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \mathrm{Col}_1(A) = \begin{bmatrix} \cos\phi \\ \sin\phi \end{bmatrix}$

$L_A \begin{bmatrix} 0 \\ 1 \end{bmatrix} = \mathrm{Col}_2(A) = \begin{bmatrix} -\sin\phi \\ \cos\phi \end{bmatrix}$

$= \begin{bmatrix} \cos(\phi + \pi/2) \\ \sin(\phi + \pi/2) \end{bmatrix}$

$\{L_A(e_1), L_A(e_2)\}$ is another o.n. basis of $\mathbb{R}^2$, just $S = \{e_1, e_2\}$ rotated (c.c.w) by angle ϕ. This L_A preserves lengths and angles

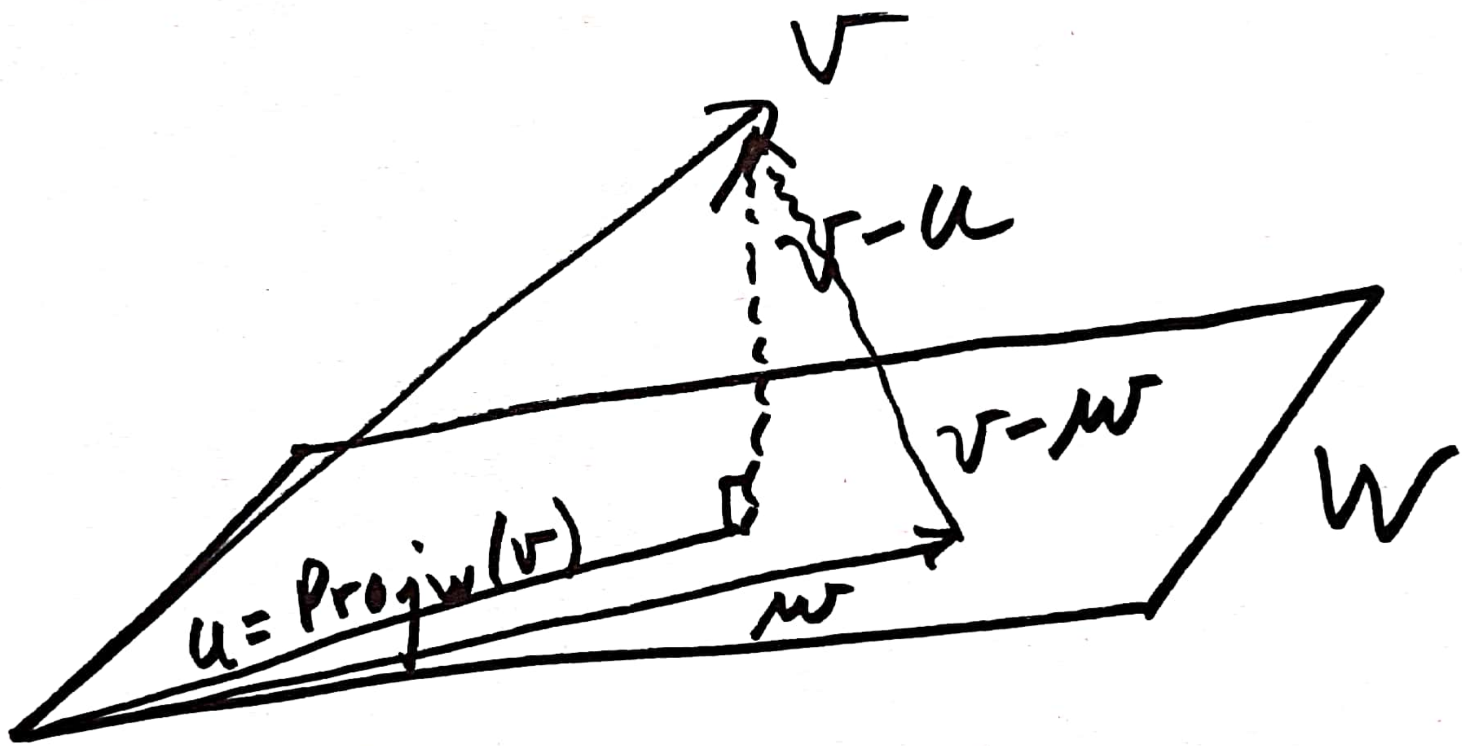
Ex: Reflections: $L_A : \mathbb{R}^2 \longrightarrow \mathbb{R}^2$ is refl. w.r.t. $y = x$, so $L_A \begin{bmatrix} x \\ y \end{bmatrix} = \begin{bmatrix} y \\ x \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} x \\ y \end{bmatrix}$

$A = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$ has columns $\left\{ \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 1 \\ 0 \end{bmatrix} \right\}$ an o.n. basis of $\mathbb{R}^2$

$A^T = A^{-1}$ so A is orthog. matrix.

L_A preserved lengths & angles.

Meaning of $\mathrm{Proj}_W(v)$ as "best approximation to v in W"



$$\|v-u\| < \|v-w\| \quad \text{if } w \neq u \quad \forall w \in W$$

$\leq$

"Best approx. Thm"