DS-GA 1014 Optimization and Computational Linear Algebra

Lab 5: Orthogonal Matrices & Eigenvalues

prepared by Ying Wang

Gram-Schmidt

Let $v_1, \ldots, v_m \in \mathbb{R}^n$ be linearly independent. Show there is an orthonormal basis for $\text{Span}(v_1, \ldots, v_m)$.

Solution. We will outline an algorithm known as Gram-Schmidt.

(a) Set $w_1 = v_1$ and $u_1 = w_1/\|w_1\|$.

(b) For $i = 2, \ldots, m$:

   i. Define w_i by
$$w_i = v_i - \langle v_i, u_1 \rangle u_1 - \cdots - \langle v_i, u_{i-1} \rangle u_{i-1}$$

   ii. Let $u_i = w_i/\|w_i\|$.

We claim that $u_1, \ldots, u_m$ are orthonormal and that $u_i \in \text{Span}(v_1, \ldots, v_i)$ for all i. The claim implies $\text{Span}(u_1, \ldots, u_i) \subseteq \text{Span}(v_1, \ldots, v_i)$ with both spans having dimension i for all $i = 1, \ldots, m$. This shows the spans are equal and completes the proof.

Previous lab question from Brett Bernstein

Gram-Schmidt

Proof of claim. Proof by induction. More precisely, we show that for all $i \geq 1$ we have $\langle u_i, u_j \rangle = 0$ for any $j < i$, $\langle u_i, u_i \rangle = 1$, and $u_i \in \mathrm{Span}(v_1, \ldots, v_i)$. For the base case $i = 1$ we only need that $v_1 \neq 0$ (so that u_1 is well-defined), but this is immediate from linear independence. For the induction case, assume the statement holds up to $i \geq 1$. By the definition of w_{i+1} and the induction hypothesis we have

$$w_{i+1} \in \mathrm{Span}(v_{i+1}, u_1, \ldots, u_i) \subseteq \mathrm{Span}(v_{i+1}, v_1, \ldots, v_i).$$

If $w_{i+1} = 0$ then $v_{i+1} \in \mathrm{Span}(v_1, \ldots, v_i)$ contradicting linear independence. Thus $w_{i+1} \neq 0$, u_{i+1} is well-defined, and $\|u_{i+1}\| = 1$. Since $u_{i+1} = w_{i+1}/\|w_{i+1}\|$ we also have

$$u_{i+1} \in \mathrm{Span}(v_{i+1}, v_1, \ldots, v_i).$$

Furthermore, for any $j < i + 1$ we have

$$\begin{aligned}
\|w_{i+1}\| \langle u_{i+1}, u_j \rangle &= \langle w_{i+1}, u_j \rangle \\
&= \langle v_{i+1} - \textstyle\sum_{k=1}^{i} \langle v_{i+1}, u_k \rangle u_k, u_j \rangle \\
&= \langle v_{i+1}, u_j \rangle - \textstyle\sum_{k=1}^{i} \langle v_{i+1}, u_k \rangle \langle u_k, u_j \rangle \\
&= \langle v_{i+1}, u_j \rangle - \langle v_{i+1}, u_j \rangle \qquad \text{(Induction Hypothesis)} \\
&= 0.
\end{aligned}$$

Gram-Schmidt

What is the output of Gram-Schmidt if the input vectors $v_1, \ldots, v_m$ are already orthonormal?

It simply sets $u_i = v_i$.

Previous lab question from Brett Bernstein

Gram-Schmidt

Let $V = R^3$ with the Euclidean inner product. We will apply the Gram-Schmidt algorithm to orthogonalize the basis $\{(1, -1, 1), (1, 0, 1), (1, 1, 2)\}$.

$\|v_1\| = \sqrt{1+1+1}$

$= \sqrt{3}$

$\|v_2\| = \sqrt{\frac{1}{9}+\frac{4}{9}+\frac{1}{9}}$

$= \sqrt{\frac{2}{3}}$

$\|v_3\| = \sqrt{\frac{1}{4}+0+\frac{1}{4}}$

$= \sqrt{\frac{1}{2}}$

Step 1 $\mathbf{v}_1 = (1, -1, 1)$.

$u_1 = \frac{1}{\sqrt{3}}(1,-1,1)$

Step 2

$$\begin{aligned} \mathbf{v}_2 &= (1,0,1) - \frac{(1,0,1)\cdot(1,-1,1)}{\|(1,-1,1)\|^2}(1,-1,1) \\ &= (1,0,1) - \tfrac{2}{3}(1,-1,1) \\ &= (\tfrac{1}{3}, \tfrac{2}{3}, \tfrac{1}{3}). \end{aligned}$$

$u_2 = \dfrac{(\frac{1}{3}, \frac{2}{3}, \frac{1}{3})}{\sqrt{\frac{2}{3}}} = \frac{\sqrt{6}}{2}(\frac{1}{3}, \frac{2}{3}, \frac{1}{3})$

Step 3

$1-1+2=2$

$\frac{1}{3}+\frac{2}{3}+\frac{2}{3}=\frac{5}{6}$

$$\begin{aligned} \mathbf{v}_3 &= (1,1,2) - \frac{(1,1,2)\cdot(1,-1,1)}{\|(1,-1,1)\|^2}(1,-1,1) - \frac{(1,1,2)\cdot(\frac{1}{3},\frac{2}{3},\frac{1}{3})}{\|(\frac{1}{3},\frac{2}{3},\frac{1}{3})\|^2}(\tfrac{1}{3}, \tfrac{2}{3}, \tfrac{1}{3}) \\ &= (1,1,2) - \tfrac{2}{3}(1,-1,1) - \tfrac{5}{2}(\tfrac{1}{3}, \tfrac{2}{3}, \tfrac{1}{3}) \\ &= (\tfrac{-1}{2}, 0, \tfrac{1}{2}). \end{aligned}$$

$u_3 = \dfrac{(-\frac{1}{2}, 0, \frac{1}{2})}{\sqrt{\frac{1}{2}}} = \sqrt{2}(-\frac{1}{2}, 0, \frac{1}{2})$

You can verify that $\{(1,-1,1), (\frac{1}{3}, \frac{2}{3}, \frac{1}{3}), (\frac{-1}{2}, 0, \frac{1}{2})\}$ forms an orthogonal basis for R^3. Normalizing the vectors in the orthogonal basis, we obtain the orthonormal basis

$$\left\{ \left(\frac{\sqrt{3}}{3}, \frac{-\sqrt{3}}{3}, \frac{\sqrt{3}}{3}\right), \left(\frac{\sqrt{6}}{6}, \frac{\sqrt{6}}{3}, \frac{\sqrt{6}}{6}\right), \left(\frac{-\sqrt{2}}{2}, 0, \frac{\sqrt{2}}{2}\right) \right\}.$$

Orthonormal bases

We can also use coordinatization for $\mathbb{R}^n$. If we have a basis $B = v_1, \ldots, v_n$ for $\mathbb{R}^n$ then we can define the coordinatization (or change-of-basis) map $\Phi_B : \mathbb{R}^n \to \mathbb{R}^n$ by

$$\Phi_B(\alpha_1 v_1 + \cdots + \alpha_n v_n) = \alpha.$$

(a) Let B denote the basis $(1,0), (-1,1)$ for $\mathbb{R}^2$. Compute

$$\Phi_B\left(\begin{bmatrix}1\\0\end{bmatrix}\right), \quad \Phi_B\left(\begin{bmatrix}-1\\1\end{bmatrix}\right), \quad \text{and} \quad \Phi_B\left(\begin{bmatrix}0\\1\end{bmatrix}\right).$$

(b) Suppose $B = v_1, \ldots, v_n$ is a basis for $\mathbb{R}^n$. Give the matrices corresponding to Φ_B and Φ_B^{-1} (possible since $\Phi_B : \mathbb{R}^n \to \mathbb{R}^n$ is linear and invertible).

(c) For which bases B of $\mathbb{R}^n$ does Φ_B preserve inner products? That is, for which bases B does

$$\langle \Phi_B(x), \Phi_B(y) \rangle = \langle x, y \rangle$$

for all $x, y \in \mathbb{R}^n$?

Previous lab question from Brett Bernstein

Orthonormal bases

$$\begin{bmatrix}1\\0\end{bmatrix} = \alpha_1 \begin{bmatrix}1\\0\end{bmatrix} + \alpha_2 \begin{bmatrix}-1\\1\end{bmatrix}$$

$\Rightarrow \alpha_1 = 1 \quad \alpha_2 = 0$

$$\begin{bmatrix}-1\\1\end{bmatrix} = \alpha_1 \begin{bmatrix}1\\0\end{bmatrix} + \alpha_2 \begin{bmatrix}-1\\1\end{bmatrix}$$

$\Rightarrow \alpha_1 = 0 \quad \alpha_2 = 1$

$$\begin{bmatrix}1\\1\end{bmatrix} = \alpha_1 \begin{bmatrix}0\\1\end{bmatrix} + \alpha_2 \begin{bmatrix}1\\1\end{bmatrix}$$

$\Rightarrow \alpha_1 = 1 \quad \alpha_2 = 1$

(a)

$$\Phi_B\left(\begin{bmatrix}1\\0\end{bmatrix}\right) = \begin{bmatrix}1\\0\end{bmatrix},$$

$$\Phi_B\left(\begin{bmatrix}-1\\1\end{bmatrix}\right) = \begin{bmatrix}0\\1\end{bmatrix},$$

$$\Phi_B\left(\begin{bmatrix}0\\1\end{bmatrix}\right) = \begin{bmatrix}1\\1\end{bmatrix}.$$

(b) Let $A \in \mathbb{R}^{n\times n}$ denote the matrix with v_i as its ith column. Then $\Phi_B = A^{-1}$ and $\Phi_B^{-1} = A$.

(c) Orthonormal bases

$$\langle Qx, Qy\rangle = (Qx)^T(Qy) = x^TQ^TQy = x^Ty = \langle x, y\rangle.$$

♡ intuition ($\mathbb{R}^2$) ♡

$\Phi_B\, v_1 = \begin{bmatrix}1\\0\end{bmatrix}$

$\Phi_B\, v_2 = \begin{bmatrix}0\\1\end{bmatrix}$

$\Rightarrow \underbrace{\Phi_B}_{A^{-1}} \underbrace{[v_1 \; v_2]}_{A} = \begin{bmatrix}1 & 0\\0 & 1\end{bmatrix} = I_2$

Previous lab question from Brett Bernstein

Recall:

[special case]

A orthogonal ($AA^T = I$)

$\Phi = A^T$

Eigenvalues & Eigenvectors

Consider the matrix

$$A=\begin{pmatrix}2 & 2\\ -4 & 8\end{pmatrix} \quad \text{and vectors} \quad v=\begin{pmatrix}1\\1\end{pmatrix} \qquad w=\begin{pmatrix}2\\1\end{pmatrix}.$$

Which are eigenvectors? What are their eigenvalues?

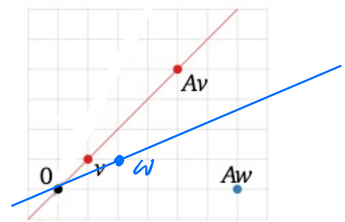
We have

$$Av=\begin{pmatrix}2 & 2\\ -4 & 8\end{pmatrix}\begin{pmatrix}1\\1\end{pmatrix}=\begin{pmatrix}4\\4\end{pmatrix}=4v.$$

Hence, v is an eigenvector of A, with eigenvalue $\lambda=4$. On the other hand,

$$Aw=\begin{pmatrix}2 & 2\\ -4 & 8\end{pmatrix}\begin{pmatrix}2\\1\end{pmatrix}=\begin{pmatrix}6\\0\end{pmatrix}.$$

which is not a scalar multiple of w. Hence, w is not an eigenvector of A.



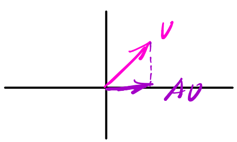
- How about $A=\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}$?

Note it's actually the x-axis projection

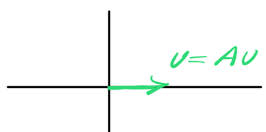
eigenvalue: 1

[verify]

$\begin{bmatrix}1 & 0\\ 0 & 0\end{bmatrix}\begin{bmatrix}\alpha\\ 0\end{bmatrix}=\begin{bmatrix}\alpha\\ 0\end{bmatrix}$ for $\forall \alpha\in\mathbb{R}$.



× (Av, v not in the same line)



✓ (Av, v in the same line)

Eigenvalues & Eigenvectors

Let

$$A = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \qquad v = \begin{pmatrix} -3 \\ 1 \end{pmatrix}.$$

Is v an eigenvector of A? If so, what is its eigenvalue?

The product is

$$Av = \begin{pmatrix} 1 & 3 \\ 2 & 6 \end{pmatrix} \begin{pmatrix} -3 \\ 1 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0v.$$

Hence, v is an eigenvector with eigenvalue *zero*.

As noted above, an eigen*value* is allowed to be zero, but an eigen*vector* is not.

Note that if $\ker(A) \neq \{0\}$

for $\forall$ nonzero $v \in \ker(A)$

$Av = 0 = 0v$

$\Rightarrow$ v is an eigenvector of A