Math 22 –
Linear Algebra and its applications

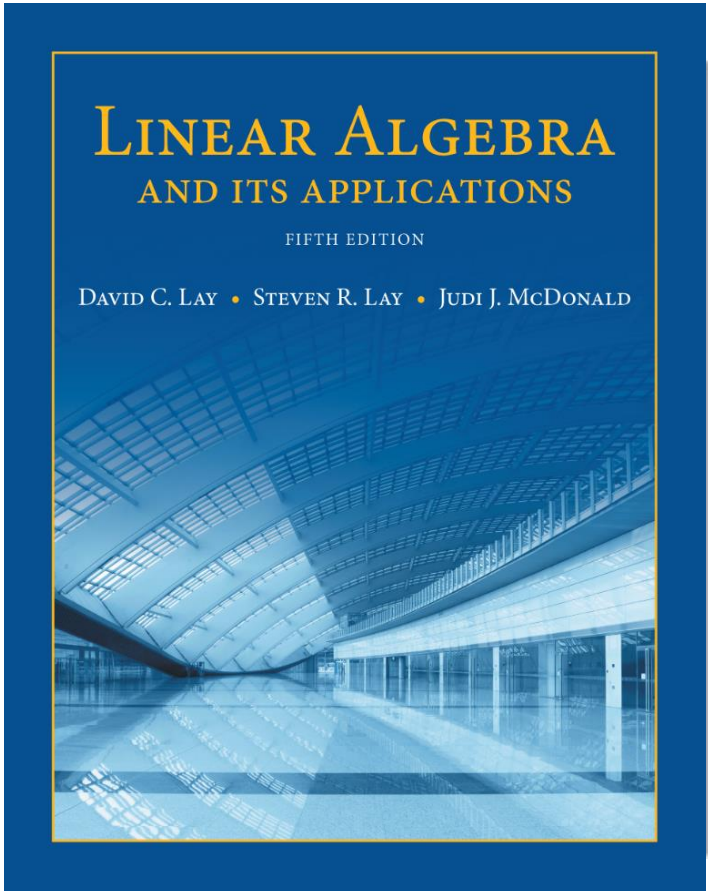
- Lecture 23 -

Instructor: Bjoern Muetzel

6 Orthogonality and Least Squares

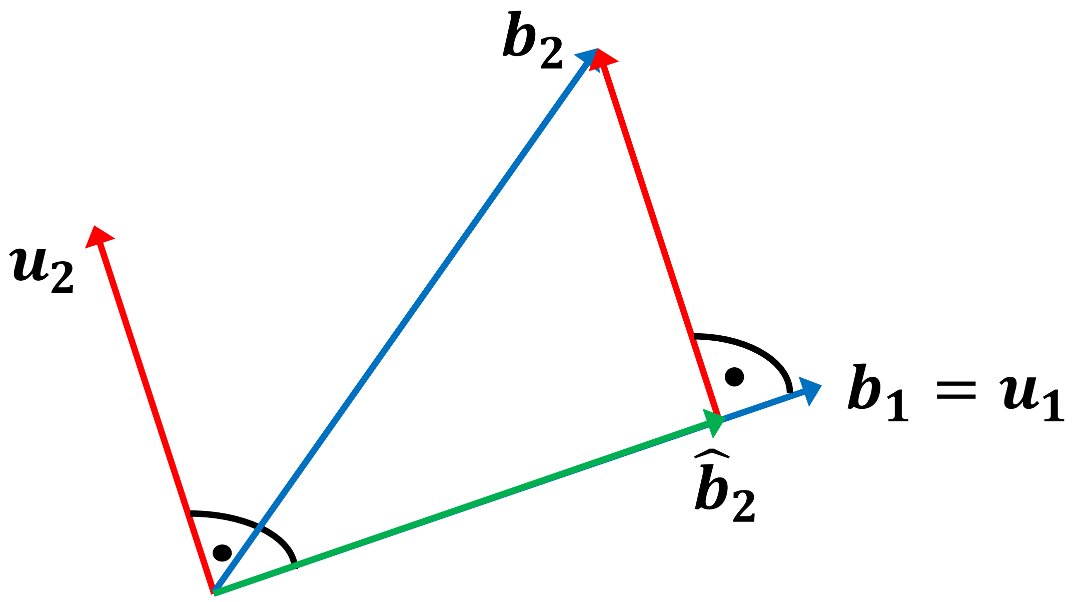
6.4

THE GRAM-SCHMIDT PROCESS



Summary:

If B={$b_1, \ldots, b_p$} is a basis of a subspace *W*. Then we can find an **orthogonal basis** for *W*. The idea is to **project** $\boldsymbol{b}_k$ **orthogonally** onto the **subspace spanned by** the previous vectors **{**$\boldsymbol{b}_1, \ldots, \boldsymbol{b}_{k-1}$**}**.



GEOMETRIC INTERPRETATION

THE GRAM-SCHMIDT PROCESS

- **Theorem 11: (Gram-Schmidt Process)**

  Given a basis $\{b_1, \ldots, b_p\}$ for a nonzero subspace W of $\mathbb{R}^n$, define

$$u_1 = b_1$$

$$u_2 = b_2 - \frac{b_2 \cdot u_1}{u_1 \cdot u_1} u_1 \qquad (1)$$

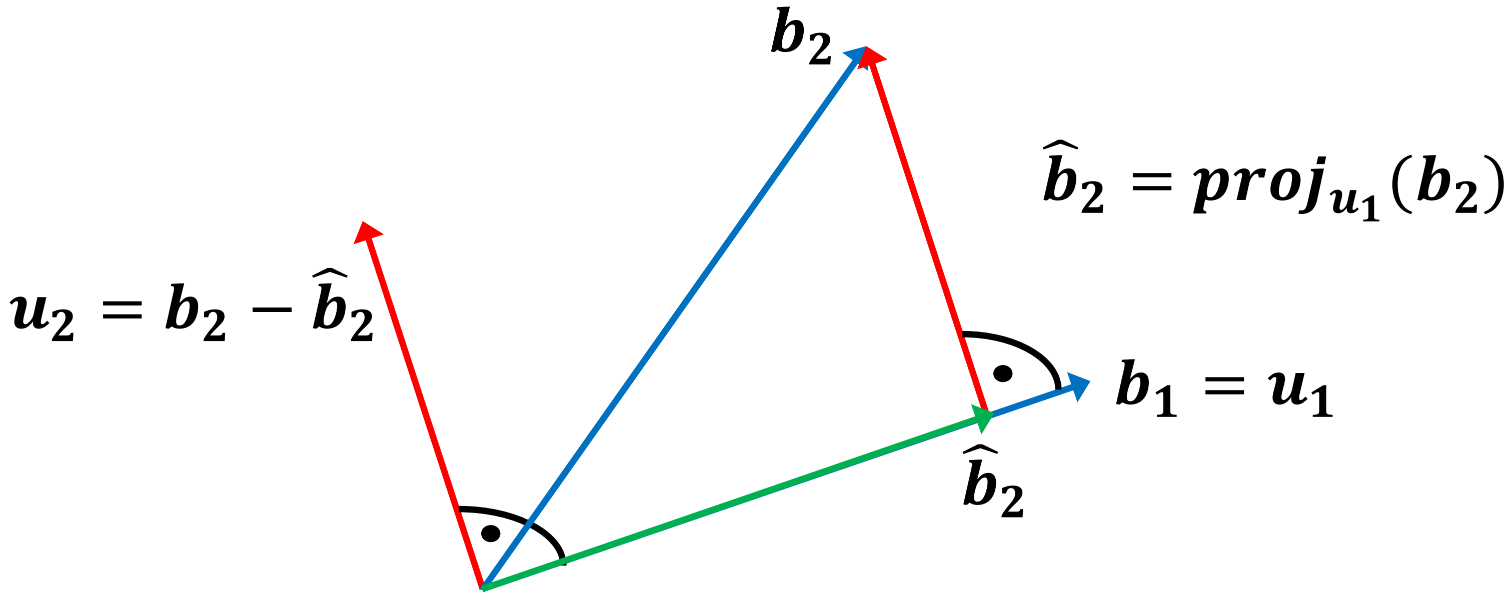
$$u_3 = b_3 - \frac{b_3 \cdot u_1}{u_1 \cdot u_1} u_1 - \frac{b_3 \cdot u_2}{u_2 \cdot u_2} u_2$$

$$\vdots$$

$$u_p = b_p - \frac{b_p \cdot u_1}{u_1 \cdot u_1} u_1 - \frac{b_p \cdot u_2}{u_2 \cdot u_2} u_1 - \ldots - \frac{b_p \cdot u_{p-1}}{u_{p-1} \cdot u_{p-1}} u_{p-1}.$$

Then $\{u_1, \ldots, u_p\}$ is an orthogonal basis for W. In addition

$$\mathbf{Span}\{\mathbf{u}_1, \ldots, \mathbf{u}_k\} = \mathbf{Span}\{\mathbf{b}_1, \ldots, \mathbf{b}_k\} \text{ for } 1 \leq k \leq p.$$



Theorem 11*: (Gram-Schmidt Process) Given a basis $\{b_1,\ldots, b_p\}$ for a nonzero subspace W of $\mathbb{R}^n$, define $\boldsymbol{W_k = \text{Span}\{b_1, \ldots, b_k\}}$ and

$$u_1 = b_1 \qquad \text{is in } W_1$$

$$u_2 = b_2 - proj_{W_1}(b_2) \qquad \text{is in } W_2$$

$$u_3 = b_3 - proj_{W_2}(b_3) \qquad \text{is in } W_3$$

$$\vdots$$

$$u_p = b_p - proj_{W_{p-1}}(b_p) \qquad \text{is in } W_p = W.$$

Then $\{u_1,\ldots, u_p\}$ is an orthogonal basis for W and $\text{Span}\{u_1,\ldots, u_k\} = W_k$.

THE GRAM-SCHMIDT PROCESS

Proof Recall that $\boldsymbol{W_k} = \textbf{Span}\{\mathbf{b_1}, \ldots, \mathbf{b_k}\}$.

1.) Set $u_1 = b_1$, then $\text{Span}\{u_1\} = \text{Span}\{b_1\}$.

2.) Suppose, for some $k < p$, we have constructed $u_1, \ldots, u_k$ so that $\{u_1, \ldots, u_k\}$ is an orthogonal basis for W_k. We set

$$\mathbf{u_{k+1} = b_{k+1} - proj_{W_k} b_{k+1}}$$

Then i.) u_{k+1} is in $W_k^{\perp}$ by the **Orthogonal Decomp. Theorem**.

ii.) $u_{k+1} \neq 0$ as b_{k+1} is not in $W_k = \text{Span}\{b_1, \ldots, b_k\}$.

Hence $\{u_1, \ldots, u_{k+1}\}$ is an orthogonal set of nonzero vectors in W_{k+1} and $\dim(W_{k+1}) = \mathbf{k+1}$.

By the **Basis Theorem** in **Sect. 4.5**, this set is a basis for W_{k+1}.

Hence $W_{k+1} = \text{Span}\{u_1, \ldots, u_{k+1}\}$.

3.) When $k + 1 = p$, the process stops.

Note: Theorem 5, **8** and **11** all make use of the same formula for orthogonal projection $\hat{\boldsymbol{y}} = \boldsymbol{proj_W}(\boldsymbol{y})$ of a vector y onto a subspace W=Span{u_1 ,…,u_p}, where {u_1,…, u_p} is an orthogonal basis:

$$\hat{\boldsymbol{y}} = \frac{\boldsymbol{y}\cdot\boldsymbol{u_1}}{\boldsymbol{u_1}\cdot\boldsymbol{u_1}}\boldsymbol{u_1} + \cdots + \frac{\boldsymbol{y}\cdot\boldsymbol{u_p}}{\boldsymbol{u_p}\cdot\boldsymbol{u_p}}\boldsymbol{u_p}.$$

This formula can be easily remembered by noticing that due to the orthogonality

$$\hat{\boldsymbol{y}} = \boldsymbol{proj_{u_1}}(\boldsymbol{y}) + \cdots + \boldsymbol{proj_{u_p}}(\boldsymbol{y}).$$

THE GRAM-SCHMIDT PROCESS

- **Example 1:** Let $b_1 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \end{bmatrix}$, $b_2 = \begin{bmatrix} 1 \\ 1 \\ 0 \\ 0 \end{bmatrix}$, $b_3 = \begin{bmatrix} 0 \\ 3 \\ 1 \\ 2 \end{bmatrix}$.

1.) Find an orthogonal basis $\{u_1, u_2\}$ for Span$\{b_1, b_2\}$.

2.) Complete $\{u_1, u_2\}$ to an to an orthogonal basis for Span$\{b_1, b_2, b_3\}$.

3.) What do we have to do to get an orthonormal basis?

QR FACTORIZATION

Theorem 12: (QR Factorization) If A is an $m \times n$ matrix with **linearly independent columns**, then A can be factored as

$A = QR$, where

i.) Q is an $m \times n$ matrix whose columns form an **orthonormal basis** for Col A.

ii.) R is an $n \times n$ **upper triangular invertible** matrix with positive entries on its diagonal.

- **Proof** The columns of A form a basis $\{a_1, \ldots, a_n\}$ for Col A.

  1.) Construct an orthogonal basis $\{u_1, \ldots, u_n\}$ for W = Col A as in **Theorem 11** (see below). Using matrix notation we get:

$$\boldsymbol{A} = [a_1, \ldots, a_n] = [u_1, \ldots, u_n]\tilde{R} = \boldsymbol{U}\tilde{\boldsymbol{R}}. \quad (\tilde{R} \text{ upper triangular})$$

$$u_1 = a_1$$
$$u_2 = a_2 - \frac{a_2 \cdot u_1}{u_1 \cdot u_1} u_1$$
$$u_3 = a_3 - \frac{a_3 \cdot u_1}{u_1 \cdot u_1} u_1 - \frac{a_3 \cdot u_2}{u_2 \cdot u_2} u_2 \qquad \text{(rearrange)}$$
$$\vdots$$
$$u_n = a_n - \frac{a_n \cdot u_1}{u_1 \cdot u_1} u_1 - \frac{a_n \cdot u_2}{u_2 \cdot u_2} u_1 - \ldots - \frac{a_n \cdot u_{n-1}}{u_{n-1} \cdot u_{n-1}} u_{n-1}$$

2.) We scale $\{u_1, \ldots, u_n\}$ to get an orthonormal basis $\{q_1, \ldots, q_n\}$. In matrix notation this translates to

$$\boldsymbol{UD} = \boldsymbol{Q} = [\mathbf{q}_1, \ldots, \mathbf{q}_n],$$ where $\boldsymbol{D}$ is a diagonal matrix.

Hence $$A = UDD^{-1}\tilde{R} = (UD)(D^{-1}\tilde{R}) = QR.$$

How can we find Q and R?

We can find $\boldsymbol{Q}$ by finding $\boldsymbol{U}$ using **Theorem 11** and then normalizing the orthogonal basis. We can obtain $\boldsymbol{R}$ by recalling that $Q^T Q = I_n$. Hence $A = QR$ implies that

$$\boldsymbol{Q}^T\mathbf{A} = \mathbf{R}.$$

Example 2: Let $A = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 2 \\ 0 & 1 & 0 \end{bmatrix}$. Recall that the columns of A are the vectors from **Example 1**. An orthogonal basis for Col$A$ = Span$\{u_1, u_2, u_3\}$ was found in that example. It was

$$[u_1, u_2, u_3] = U = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 2 \\ 0 & 1 & -1 \end{bmatrix}.$$

Find Q and R, such that $A = QR$ as in **Theorem 12**.