E.g. 7.3.13. *Let* $\mathbb{R}^2$ *have the Euclidean inner product. Use the Gram-Schmidt algorithm to construct an orthonormal basis from* $\{\mathbf{u}_1, \mathbf{u}_2\}$ *where* $\mathbf{u}_1 = (1, -3)$ *and* $\mathbf{u}_2 = (2, 2)$.

- $\underset{\sim}{v}_1 = \underset{\sim}{u}_1$

- $\underset{\sim}{u}_2 - \text{proj}_{W_1} u_2$ where $W_1 = \text{span}\{\underset{\sim}{v}_1\}$

$$= \underset{\sim}{u}_2 - \frac{\langle \underset{\sim}{u}_2, \underset{\sim}{v}_1 \rangle}{\langle \underset{\sim}{v}_1, \underset{\sim}{v}_1 \rangle} \underset{\sim}{v}_1$$

$$= (2,2) - \frac{(2,2)\cdot(1,-3)}{(1,-3)\cdot(1,-3)}(1,-3) = \frac{4}{5}(3,1)$$

So choose:

$\underset{\sim}{v}_2 = (3,1)$

- Now normalise:

$$\underset{\sim}{w}_1 = \frac{\underset{\sim}{v}_1}{\|\underset{\sim}{v}_1\|} = \frac{1}{\sqrt{10}}(1,-3)$$

$$\underset{\sim}{w}_2 = \frac{\underset{\sim}{v}_2}{\|\underset{\sim}{v}_2\|} = \frac{1}{\sqrt{10}}(3,1).$$

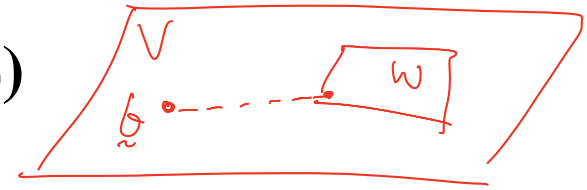
Theorem 7.3.14. *If V is a finite-dimensional inner product space, and W is a nonzero subspace of V, then $V = W \oplus W^{\perp}$.*

Proof. Since V is finite dimensional, so is W.
Since W is also nonzero, Th^{m} 7.3.12 $\Rightarrow$ W has an orthonormal basis. Th^{m} 7.3.8 $\Rightarrow$ every $\underset{\sim}{u} \in V$ can be written as $\underset{\sim}{u} = \underset{\sim}{w} + \underset{\sim}{w}_{\perp}$ where $\underset{\sim}{w} \in W$ & $\underset{\sim}{w}_{\perp} \in W^{\perp}$ & so $V = W + W^{\perp}$

Since $W^{\perp} \perp W$ by def^{n}, Th^{m} 7.2.9 $\Rightarrow$ $W^{\perp} \cap W = \{\underset{\sim}{0}\}$ and so the sum $W + W^{\perp}$ is direct.

☐

7.4 Least Squares (A&R §6.4)



We start with a geometric problem: Suppose W is a subspace of an inner product space V, and $\mathbf{b} \in V$. How do we find the vector in W which is closest to $\mathbf{b}$?

Theorem 7.4.1 (Closest Point)**.** *If W is a finite dimensional subspace of an inner product space V and $\mathbf{b} \in V$, then the point in W closest to $\mathbf{b}$ is $\mathrm{proj}_W\mathbf{b}$, in the sense that*

$$d(\mathbf{b}, \mathrm{proj}_W\mathbf{b}) < d(\mathbf{b}, \mathbf{w})$$

for every vector $\mathbf{w} \in W$ different from $\mathrm{proj}_W\mathbf{b}$.

Proof. For any $\underset{\sim}{w} \in W$

$\underset{\sim}{b} - \underset{\sim}{w} = (\underset{\sim}{b} - \mathrm{proj}_W \underset{\sim}{b}) + (\mathrm{proj}_W \underset{\sim}{b} - \underset{\sim}{w})$

with $\mathrm{proj}_W \underset{\sim}{b} - \underset{\sim}{w} \in W$ & $\underset{\sim}{b} - \mathrm{proj}_W \underset{\sim}{b} \in W^\perp$

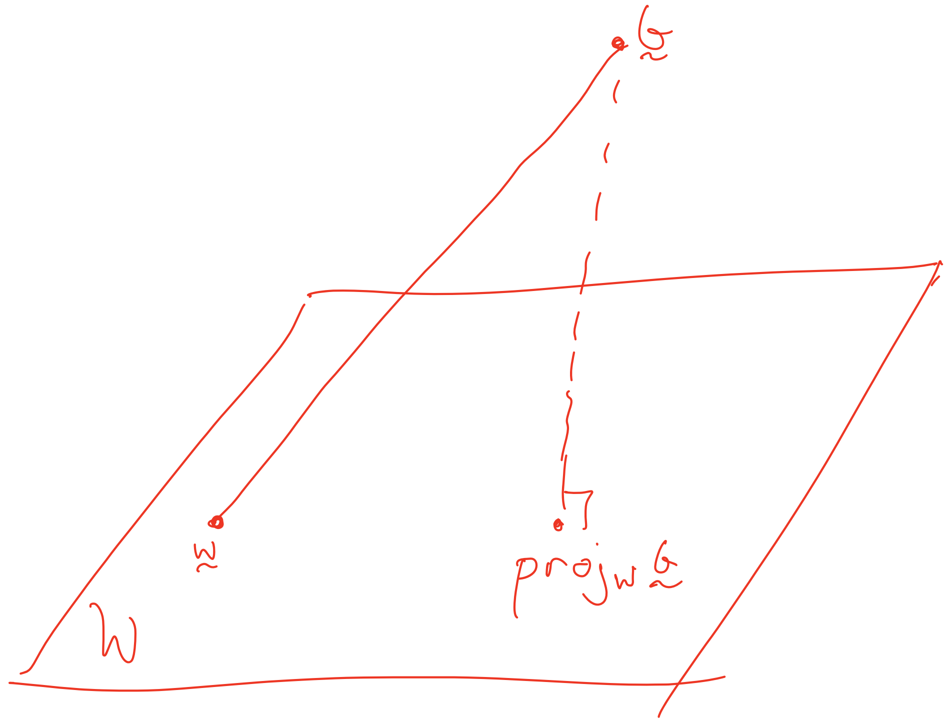
So Pythagoras $\Rightarrow$ $\|\underset{\sim}{b} - \underset{\sim}{w}\|^2 = \|\underset{\sim}{b} - \mathrm{proj}_W \underset{\sim}{b}\|^2 + \|\mathrm{proj}_W \underset{\sim}{b} - \underset{\sim}{w}\|^2$

If $\underset{\sim}{w} \neq \mathrm{proj}_W \underset{\sim}{b}$ then $\|\mathrm{proj}_W \underset{\sim}{b} - \underset{\sim}{w}\|^2 > 0$ by positive definiteness $\Rightarrow$

$\|\underset{\sim}{b} - \underset{\sim}{w}\|^2 > \|\underset{\sim}{b} - \mathrm{proj}_W \underset{\sim}{b}\|^2 \Rightarrow$

$\|\underset{\sim}{b} - \underset{\sim}{w}\| > \|\underset{\sim}{b} - \mathrm{proj}_W \underset{\sim}{b}\|$

$d(\underset{\sim}{b}, w) > d(\underset{\sim}{b}, \mathrm{proj}_W \underset{\sim}{b})$ $\square$



b
w
proj_W b
W

□

In addition to being of geometric interest, Theorem 7.4.1 has a practical application to the least-squares problem.

Problem 7.4.2. *Given a linear system $A\mathbf{x} = \mathbf{b}$ of m equations in n unknowns, find a vector $\mathbf{x}$ that minimizes $\|\mathbf{b} - A\,\mathbf{x}\|$ with respect to the Euclidean inner product on $\mathbb{R}^m$. We call such an $\mathbf{x}$ a **least squares solution** of the system, we call $\mathbf{b} - A\mathbf{x}$ the **least squares error vector** and $\|\mathbf{b} - A\mathbf{x}\|$ the **least squares error**.*

Remark 7.4.3. *The term "least squares solution" arises because the error vector $\mathbf{b} - A\mathbf{x} = (e_1, e_2, \ldots, e_m)$, has squared length $\|\mathbf{b} - A\mathbf{x}\|^2 = e_1^2 + e_2^2 + \ldots + e_m^2$. Since minimizing $\|\mathbf{b} - A\mathbf{x}\|$ is equivalent to minimizing $\|\mathbf{b} - A\mathbf{x}\|^2$, the least squares solution, as defined above, minimizes the "sum of the squares of the errors".*

Theorem 7.4.4. *Let $W = \mathrm{col}(A)$. Then $\mathbf{x}$ is a least squares solution of $A\mathbf{x} = \mathbf{b}$ iff $\boxed{A\mathbf{x} = \mathrm{proj}_W\mathbf{b}}$.*

Proof. By def^n, $\underset{\sim}{x} \in \mathbb{R}^n$ is a least squares sol^n of $A\underset{\sim}{x} = \underset{\sim}{b}$ iff $\underset{\sim}{x}$ minimizes $d(\underset{\sim}{b}, A\underset{\sim}{x})$. For any $\underset{\sim}{x} \in \mathbb{R}^n$, $A\underset{\sim}{x} \in \mathrm{col}(A)$. Th^m 7.4.1 $\Rightarrow$ $\mathrm{proj}_W \underset{\sim}{b}$ is the closest pt in W to $\underset{\sim}{b}$. So $d(\underset{\sim}{b}, A\underset{\sim}{x})$ is minimized iff $\underset{\sim}{x}$ satisfies $A\underset{\sim}{x} = \mathrm{proj}_W \underset{\sim}{b}$.

$\square$

$\square$

Theorem 7.4.5. *A vector* $\mathbf{x}$ *is a least squares solution to* $A\mathbf{x} = \mathbf{b}$ *iff it is a solution of the associated* ***normal system***

$$A^T A\mathbf{x} = A^T\mathbf{b}. \tag{7.5}$$

Moreover, the normal system is always consistent.

Proof. Th^{m} 7.4.4 $\Rightarrow$ $\underset{\sim}{x}$ is a least squares sol^{n} of $A\underset{\sim}{x} = \underset{\sim}{b}$ iff $A\underset{\sim}{x} = \text{proj}_{col(A)}\underset{\sim}{b}$.

- If $A\underset{\sim}{x} = \text{proj}_{col(A)}\underset{\sim}{b}$ then

  $\underset{\sim}{b} - A\underset{\sim}{x} = \underset{\sim}{b} - \text{proj}_{col(A)}\underset{\sim}{b} \in col(A)^{\perp} = null(A^T)$

  $\Rightarrow \quad A^T(\underset{\sim}{b} - A\underset{\sim}{x}) = \underset{\sim}{0} \Rightarrow A^TA\underset{\sim}{x} = A^T\underset{\sim}{b}$

  In particular, since $\text{proj}_{col(A)}\underset{\sim}{b} \in col(A)$

  the eq^{n} $A\underset{\sim}{x} = \text{proj}_{col(A)}\underset{\sim}{b}$ is consistent $\Rightarrow$

  $A^TA\underset{\sim}{x} = A^T\underset{\sim}{b}$ is consistent.

- Conversely if $A^TA\underset{\sim}{x} = A^T\underset{\sim}{b}$ then

  $A^T(A\underset{\sim}{x} - \underset{\sim}{b}) = \underset{\sim}{0} \Leftrightarrow A\underset{\sim}{x} - \underset{\sim}{b} \in null(A^T) = col(A)^{\perp}$

  But $A\underset{\sim}{x} \in col(A) \quad \forall \underset{\sim}{x} \in \mathbb{R}^n$ & $\underset{\sim}{b} = A\underset{\sim}{x} + (\underset{\sim}{b} - A\underset{\sim}{x})$

  $\Rightarrow \quad A\underset{\sim}{x} = \text{proj}_{col(A)}\underset{\sim}{b}$ by Th^{m} 7.3.7

  $\Rightarrow \underset{\sim}{x}$ is a least squares sol^{n}

$\square$

E.g. 7.4.6. *Find all least squares solutions of* $A\mathbf{x} = \mathbf{b}$ *where*

$$A = \begin{bmatrix} 1 & 3 \\ -2 & -6 \\ 3 & 9 \end{bmatrix}, \qquad \mathbf{b} = \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}.$$

Solve $A^T A \underset{\sim}{x} = A^T \underset{\sim}{b}$

$A^T A = \begin{pmatrix} 14 & 42 \\ 42 & 126 \end{pmatrix}$ & $A^T \underset{\sim}{b} = \begin{pmatrix} 4 \\ 12 \end{pmatrix}$

$$\left(\begin{array}{cc|c} 14 & 42 & 4 \\ 42 & 126 & 12 \end{array}\right) \sim \left(\begin{array}{cc|c} 14 & 42 & 4 \\ 0 & 0 & 0 \end{array}\right)$$

$$\Longrightarrow \underset{\sim}{x} = \begin{pmatrix} 2/7 \\ 0 \end{pmatrix} + \alpha \begin{pmatrix} -3 \\ 1 \end{pmatrix} \quad \forall \, \alpha \in \mathbb{R}$$

Error: $A\underset{\sim}{x} - \underset{\sim}{b} = \begin{pmatrix} 2/7 \\ -4/7 \\ 6/7 \end{pmatrix} - \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}$

$$\Longrightarrow \|A\underset{\sim}{x} - \underset{\sim}{b}\| = \sqrt{6/7}$$

So all least squares sol$^{\underline{n}}$s have same error.