The Gram-Schmidt process

Let V be a subspace of $\mathbb{R}^n$, with a basis $\{\mathbf{w}_j\}_{j=1}^k$.

Gram-Schmidt is an algorithm for building an orthogonal basis $\{\mathbf{v}_j\}_{j=1}^k$ for V.

Step 1: Set $\mathbf{v}_1 = \mathbf{w}_1$, and $V_1 = \text{span}(\mathbf{w}_1) = \text{span}(\mathbf{v}_1)$.

Step 2: Set $\mathbf{p}_2 = \dfrac{\mathbf{v}_1 \cdot \mathbf{w}_2}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$. *$\mathbf{p}_2$ is the projection of $\mathbf{w}_2$ onto V_1.*

Set $\mathbf{v}_2 = \mathbf{w}_2 - \mathbf{p}_2$, and $V_2 = \text{span}(\mathbf{w}_1, \mathbf{w}_2) = \text{span}(\mathbf{v}_1, \mathbf{v}_2)$. *$\mathbf{v}_2 \in V_1^\perp$ by construction.*

Step 3: Set $\mathbf{p}_3 = \dfrac{\mathbf{v}_1 \cdot \mathbf{w}_3}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \dfrac{\mathbf{v}_2 \cdot \mathbf{w}_3}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$. *$\mathbf{p}_3$ is the projection of $\mathbf{w}_3$ onto V_2.*

Set $\mathbf{v}_3 = \mathbf{w}_3 - \mathbf{p}_3$, and $V_3 = \text{span}(\mathbf{w}_i)_{i=1}^3 = \text{span}(\mathbf{v}_i)_{i=1}^3$. *$\mathbf{v}_3 \in V_2^\perp$ by construction.*

Question: Could we have “bad luck” and find that $\mathbf{v}_3 = \mathbf{0}$?

No, this is not possible. If $\mathbf{v}_3 = \mathbf{0}$, then $\mathbf{w}_3 = \mathbf{p}_3 \in V_2$. Since V_2 is spanned by $\{\mathbf{w}_1, \mathbf{w}_2\}$, this would mean that the set $\{\mathbf{w}_1,\ \mathbf{w}_2,\ \mathbf{w}_3\}$ is linearly dependent. This would contract the assumption that the vectors $\{\mathbf{w}_i\}$ form a basis.

The Gram-Schmidt process: Let V be a subspace of dimension k in $\mathbb{R}^n$ (e.g. $V = \mathbb{R}^n$).

Input: A basis $\{\mathbf{w}_j\}_{j=1}^k$ for V	$\Rightarrow$	**Output:** An orthogonal basis $\{\mathbf{v}_j\}_{j=1}^k$ for V

Step 1: $\mathbf{v}_1 = \mathbf{w}_1$.

Step 2: $\mathbf{p}_2 = \dfrac{\mathbf{w}_2 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1$ $\qquad \mathbf{v}_2 = \mathbf{w}_2 - \mathbf{p}_2$.

Step 3: $\mathbf{p}_3 = \dfrac{\mathbf{w}_3 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \dfrac{\mathbf{w}_3 \cdot \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2$ $\qquad \mathbf{v}_3 = \mathbf{w}_3 - \mathbf{p}_3$.

⋮

Step t: $\mathbf{p}_t = \displaystyle\sum_{i=1}^{t-1} \frac{\mathbf{w}_t \cdot \mathbf{v}_i}{\|\mathbf{v}_i\|^2}\mathbf{v}_i$ $\qquad \mathbf{v}_t = \mathbf{w}_t - \mathbf{p}_t$.

The Gram-Schmidt process: Let V be a subspace of dimension k in $\mathbb{R}^n$ (e.g. $V = \mathbb{R}^n$).

Input: A basis $\{\mathbf{w}_j\}_{j=1}^k$ for V	$\Rightarrow$	**Output:** An orthogonal basis $\{\mathbf{v}_j\}_{j=1}^k$ for V

Step 1: $\mathbf{v}_1 = \mathbf{w}_1$.

Step t: $\mathbf{p}_t = \displaystyle\sum_{i=1}^{t-1} \frac{\mathbf{w}_t \cdot \mathbf{v}_i}{\|\mathbf{v}_i\|^2}\mathbf{v}_i$ $\qquad$ $\mathbf{v}_t = \mathbf{w}_t - \mathbf{p}_t$.

Properties of the Gram-Schmidt process: The following are true for $t \in \{1, 2, \dots, k\}$:

- $\mathbf{v}_t = \mathbf{w}_t + \sum_{j=1}^{t-1} c_j \mathbf{w}_j$.
- $\text{span}(\mathbf{v}_i)_{i=1}^t = \text{span}(\mathbf{w}_i)_{i=1}^t$.
- $\mathbf{p}_t$ is the orthogonal projection of $\mathbf{w}_t$ onto $\text{span}(\mathbf{w}_i)_{i=1}^{t-1}$.
- $\mathbf{v}_t$ is orthogonal to $\text{span}(\mathbf{w}_i)_{i=1}^{t-1}$.
- $\|\mathbf{v}_t\|$ is the distance between $\mathbf{w}_t$ and $\text{span}(\mathbf{w}_i)_{i=1}^{t-1}$. *It is never zero!*

The Gram-Schmidt process with normalization of the vectors:

Let V be a subspace of dimension k in $\mathbb{R}^n$ (e.g. $V = \mathbb{R}^n$).

Input: A basis $\{\mathbf{w}_j\}_{j=1}^k$ for V $\Rightarrow$ **Output:** An orthonormal basis $\{\mathbf{v}_j\}_{j=1}^k$ for V

Step 1: $$\mathbf{v}_1 = \frac{\mathbf{w}_1}{\|\mathbf{w}_1\|}$$

Step 2: $\mathbf{p}_2 = (\mathbf{w}_2 \cdot \mathbf{v}_1)\,\mathbf{v}_1$ $$\mathbf{v}_2 = \frac{\mathbf{w}_2 - \mathbf{p}_2}{\|\mathbf{w}_2 - \mathbf{p}_2\|}$$

Step 3: $\mathbf{p}_3 = (\mathbf{w}_3 \cdot \mathbf{v}_1)\,\mathbf{v}_1 + (\mathbf{w}_3 \cdot \mathbf{v}_2)\,\mathbf{v}_2$ $$\mathbf{v}_3 = \frac{\mathbf{w}_3 - \mathbf{p}_3}{\|\mathbf{w}_3 - \mathbf{p}_3\|}$$

⋮

Step t**:** $\mathbf{p}_t = \sum_{i=1}^{t-1} (\mathbf{w}_t \cdot \mathbf{v}_i)\,\mathbf{v}_i$ $$\mathbf{v}_t = \frac{\mathbf{w}_t - \mathbf{p}_t}{\|\mathbf{w}_t - \mathbf{p}_t\|}$$

Example: Set $\mathbf{w}_1 = [1, 2, 2]$, $\mathbf{w}_2 = [-1, 0, 2]$, $\mathbf{w}_3 = [0, 0, 1]$. Orhogonalize $\{\mathbf{w}_i\}_{i=1}^{3}$!

$$\mathbf{v}_1 = \mathbf{w}_1 = [1, 2, 2]$$

$$\boldsymbol{p}_2 = \frac{\mathbf{w}_2 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 = \frac{1 \cdot (-1) + 2 \cdot 0 + 2 \cdot 2}{1^2 + 2^2 + 2^2}[1, 2, 2] = \frac{3}{9}[1, 2, 2]$$

$$\mathbf{v}_2 = \mathbf{w}_2 - \mathbf{p}_2 = [-1, 0, 2] - \frac{1}{3}[1, 2, 2] = [-4/3,\ -2/3,\ 4/3]$$

$$\boldsymbol{p}_3 = \frac{\mathbf{w}_3 \cdot \mathbf{v}_1}{\|\mathbf{v}_1\|^2}\mathbf{v}_1 + \frac{\mathbf{w}_3 \cdot \mathbf{v}_2}{\|\mathbf{v}_2\|^2}\mathbf{v}_2 = \frac{2}{9}[1, 2, 2] + \frac{4/3}{4}[-4/3,\ -2/3,\ 4/3]$$

$$\mathbf{v}_3 = \mathbf{w}_3 - \mathbf{p}_3 = [2/9,\ -2/9,\ 1/9]$$