Gram-Schmidt Orthogonalization

Szymon Rusinkiewicz
COS 302, Fall 2020
PRINCETON UNIVERSITY

Orthonormal Bases

- Orthonormal bases have all basis vectors unit-length and perpendicular
- Nice to work with: projection of arbitrary vector along basis = dot product
- Orthonormal transformations represent rotations and reflections: no scale, no shear
- For a square orthonormal matrix $\boldsymbol{R}$, we have $\boldsymbol{R}^{-1} = \boldsymbol{R}^\top$

Gram-Schmidt

- The Gram-Schmidt process takes arbitrary basis, makes it orthonormal
- Simple, intuitive "greedy" algorithm
- Not the most stable numerically — we'll see a better solution with SVD

Gram-Schmidt

- Start with existing basis $(\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots \boldsymbol{b}_n)$
- Produce orthonormal basis $(\boldsymbol{u}_1, \boldsymbol{u}_2, \ldots \boldsymbol{u}_n)$
  (Note: your textbook only produces *orthogonal*, not *orthonormal*)

- Greedy start:

$$\boldsymbol{u}_1 = \frac{\boldsymbol{b}_1}{\|\boldsymbol{b}_1\|}$$

Gram-Schmidt

- For the second basis vector, start with $\boldsymbol{b}_2$ and remove component along $\boldsymbol{u}_1$:

$$\boldsymbol{u}_2 = \frac{\boldsymbol{b}_2 - \pi_{\boldsymbol{u}_1}\boldsymbol{b}_2}{\left\|\boldsymbol{b}_2 - \pi_{\boldsymbol{u}_1}\boldsymbol{b}_2\right\|}$$

- But projection onto a unit-length vector is just a dot product:

$$\boldsymbol{u}_2 = \frac{\boldsymbol{b}_2 - (\boldsymbol{b}_2 \cdot \boldsymbol{u}_1)\,\boldsymbol{u}_1}{\left\|\boldsymbol{b}_2 - (\boldsymbol{b}_2 \cdot \boldsymbol{u}_1)\,\boldsymbol{u}_1\right\|}$$

Gram-Schmidt

- For the third basis vector, start with $\boldsymbol{b}_3$ and remove component in the span of $\boldsymbol{u}_1$ and $\boldsymbol{u}_2$:

$$\boldsymbol{u}_3 = \frac{\boldsymbol{b}_3 - \pi_{span[\boldsymbol{u}_1,\boldsymbol{u}_2]}\boldsymbol{b}_3}{\left\|\boldsymbol{b}_3 - \pi_{span[\boldsymbol{u}_1,\boldsymbol{u}_2]}\boldsymbol{b}_3\right\|}$$

- For orthonormal basis, projection into span given by dot products along basis vectors:

$$\boldsymbol{u}_3 = \frac{\boldsymbol{b}_3 - (\boldsymbol{b}_3 \cdot \boldsymbol{u}_1)\,\boldsymbol{u}_1 - (\boldsymbol{b}_3 \cdot \boldsymbol{u}_2)\,\boldsymbol{u}_2}{\left\|\boldsymbol{b}_3 - (\boldsymbol{b}_3 \cdot \boldsymbol{u}_1)\,\boldsymbol{u}_1 - (\boldsymbol{b}_3 \cdot \boldsymbol{u}_2)\,\boldsymbol{u}_2\right\|}$$

Gram-Schmidt

- And so on...

$$\boldsymbol{u}_k = \frac{\boldsymbol{b}_k - \pi_{span[\boldsymbol{u}_1 .. \boldsymbol{u}_{k-1}]}\boldsymbol{b}_k}{\left\|\boldsymbol{b}_k - \pi_{span[\boldsymbol{u}_1 .. \boldsymbol{u}_{k-1}]}\boldsymbol{b}_k\right\|}$$

$$= \frac{\boldsymbol{b}_k - \sum_{j=1}^{k-1} (\boldsymbol{b}_k \cdot \boldsymbol{u}_j)\, \boldsymbol{u}_j}{\left\|\boldsymbol{b}_k - \sum_{j=1}^{k-1} (\boldsymbol{b}_k \cdot \boldsymbol{u}_j)\, \boldsymbol{u}_j\right\|}$$

Gram-Schmidt Example

Example:

- Start with

$$\boldsymbol{b}_1 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}, \quad \boldsymbol{b}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

- Compute

$$\boldsymbol{u}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{u}_2 = \begin{bmatrix} 1 \\ 1 \end{bmatrix} - \begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$$

Gram-Schmidt Example

Example:

- If start with a different order:

$$\boldsymbol{b}_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}, \quad \boldsymbol{b}_2 = \begin{bmatrix} 2 \\ 0 \end{bmatrix}$$

- Compute

$$\boldsymbol{u}_1 = \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix}, \quad \begin{bmatrix} 2 \\ 0 \end{bmatrix} - \left(\begin{bmatrix} 2 \\ 0 \end{bmatrix} \cdot \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} \right) \begin{bmatrix} \sqrt{2}/2 \\ \sqrt{2}/2 \end{bmatrix} = \begin{bmatrix} 1 \\ -1 \end{bmatrix} \longrightarrow \boldsymbol{u}_2 = \begin{bmatrix} \sqrt{2}/2 \\ -\sqrt{2}/2 \end{bmatrix}$$