Discussion 13B

The following note is useful for this discussion: Note 2j.

1. Gram Schmidt on Complex Vectors

(a) Consider the three complex vectors

$$\vec{a}_1 = \begin{bmatrix} \mathrm{j} \\ -1 \\ 0 \end{bmatrix}, \quad \vec{a}_2 = \begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}, \quad \vec{a}_3 = \begin{bmatrix} 0 \\ \mathrm{j} \\ 1 \end{bmatrix}. \tag{1}$$

Compute an orthonormal basis from this list of vectors with complex Gram-Schmidt.
(HINT: The complex version of Gram-Schmidt is Algorithm 45 of Note 2j.)

(b) **(PRACTICE) Derive the least-squares solution for the case of a complex tall matrix of data and a tall matrix of values.** We want to find the best (complex) linear combination of the columns for predicting the observed values in a least-squares sense — we want to minimize the norm of the residual.

This can be formulated as having a feature matrix of data $D \in \mathbb{C}^{m \times n}$ where $m > n$ and measurements $\vec{y} \in \mathbb{C}^m$. In this case, feel free to assume that the columns of D are linearly independent even when we allow complex linear combinations. **First assume that the columns of D are orthonormal.** *Hint: You may find it useful to define a matrix $U = \begin{bmatrix} D & D_+ \end{bmatrix}$, obtained via Gram-Schmidt.* Recall that the procedure here resembles that used in the SVD derivation (one portion of U contains the data/info we "care about", and the remainder is there via extension, to span the space.)

(c) **(PRACTICE) Repeat the previous part without the assumption of orthonormality for the columns of D.** You can keep the assumption of linear independence.

Contributors:

- Neelesh Ramachandran.
- Nathan Lambert.
- Anant Sahai.