

The SVD is used for the stable computation of principal component analysis (PCA).
Ringnér, *Nat Biotechnol* 26, 303 (2008).

The SVD is Different Than PCA

→ **PCA assumes preprocessing of the data, which limits the data interpretation** (e.g., the SVD of a dataset can identify the probability distribution function that is sampled by the dataset with no a-priori assumptions; **PCA cannot**).

Alter & Golub, *PNAS* 103, 11828 (2006); https://alterlab.org/harmonic_oscillator/

Cadima & Jolliffe, *Pak J Statist* 25, 473 (2009);

Muralidhara, Gross, Gutell & Alter, *PLoS One* 6, e18768 (2011); <https://www.alterlab.org/rRNA/>

Bertagnolli et al., *PLoS One* 8, e78913 (2013); https://alterlab.org/GBM_metabolism/

→ **PCA identifies patterns across the columns separately from patterns across the rows; the SVD simultaneously computes the corresponding sets of patterns across the rows and columns, ensuring consistent data interpretation.**

Alter et al., *PNAS* 97, 10101 (2000); https://alterlab.org/singular_value_decomposition/

Alter et al., in *Microarrays: Optical Technologies and Informatics*. Bellingham, WA: International Society for Optics and Photonics (SPIE) (2001); <https://alterlab.org/SVD/>

Fellenberg, Hauser, Brors, Neutzner, Hoheisel & Vingron, *PNAS* 98, 10781 (2001).

→ **PCA, as it is programmed in most computational packages, is limited to classifying the data based upon the two or three patterns that capture most of the information in the data (e.g., variance in the case of column centering); the SVD maintains all data patterns, and not just for data classification.**