Advanced Topics in Algebra – Lab#6

dr Michał Jarema

Topic: PCA and SVD for dimension reduction

Task 1 (continuation)

Do the PCA analysis using eigen(cov(X)) and svd(X), where X is the centered data matrix.

Check that right singular vectors V are principal directions(eigenvectors of covariance matrix) and that singular values are related to the eigenvalues of covariance matrix via $\lambda_i = s_i^2/(\mathrm{n}-1)$. If you have doubt read this[13] again.

Identify the principal values and vectors in the output of functions PCA() from FactoMineR, and built-in prcomp(), princomp().

Select ONE task: 2 or 2a*

Task 2

Read through this useful tutorial[14] and just repeat the analysis. Do only the sections listed on whiteboard.

Task 2a*

Do task 2 for your dataset (with $d \geq 10$ features/ variables). Select such a dataset, where you expect some correlations between its features.

Task 3
Repeat the image compression by SVD presented in lab for your selected grayscale image.

Further reading:

To scale or not to scale?[15]
Remember → using option `scale=true` bases PCA on correlation matrix (all variances ale scaled to 1). It removes information about relative variation ranges of different variables. Be careful. But sometimes this information is trivially unnecessary, e.g. when the variables are in different incomparable units, like cm, km, kg.

(*)Various methods used for dimension reduction[16]

---

[13] https://stats.stackexchange.com/q/134283

[14] http://www.sthda.com/english/articles/31-principal-component-methods-in-r-practical-guide/112-pca-principal-component-analysis-essentials/

[15] https://stats.stackexchange.com/a/78

[16] https://www.analyticsvidhya.com/blog/2015/07/dimension-reduction-methods/