BS6213

Art of Principal Component Analysis

Professor Wong Limsoon

9 January 2023

This session introduces students to interesting ways to think about and use PCA.

The supplementary materials below are useful for a deeper appreciation of the examples discussed in the session. These are optional to read:

- Jolicoeur & Mosimann, "Size and shape variation in the painted turtle: A principal component analysis", *Growth*, 24:339-354, 1960. This article gives an idea of how PCA is traditionally used.
- Giuliani et al., "On the constructive role of noise in spatial systems", *Physics Letters A*, 247:47-52, 1998. This article is useful background for exercises #1 and #2.
- Goh & Wong, "Protein complex-based analysis is resistant to the obfuscating consequences of batch effects---a case study in clinical proteomics", *BMC Genomics*, 18:142, 2017. This article is useful background for exercise #3 and #4.
- Leek et al., "Tackling the widespread and critical impact of batch effects in high-throughput data", *Nature Reviews Genetics*, 11:733-739, 2010. This article is useful background on batch effects that exercise #3 and #4 are based on.
- Wang et al., "Extensive up-regulation of gene expression in cancer: The normalized use of microarray data", Molecular Biosystems, 8:818-827, 2012. This article is useful background to appreciate misused of batch effect-normalization.

Pre- & post session reports assignment

Before the session (submit by 8 January 2023):

- Read a basic intro to PCA (e.g., https://en.wikipedia.org/wiki/Principal_component_analysis)
- Submit a brief report that discusses whether there is any conceptual or technical error in Section 3.1 "Segmentation by Eigenbackground subtraction" of this paper, https://www.nuriaoliver.com/humanbehavior/icvs99.pdf.

After the session (submit by 11 January 2023):

- Submit a brief report on what you have learned from the session. Any change in your opinion & comments on the paper from the pre-session report?