

Assignment 2

Instructor: Jiguo Cao

Instructions: It is due at 3:30pm, Tuesday, June 19, 2018.

1. Rangel *et al.* (2004) collected time course microarray data for T-cell activation to investigate a dynamic gene regulatory network. For the time course measurement of each gene, 10 unequally spaced observations are sampled; each gene has 34 independent replicates. Here we analyze all 34 replicates of the time course measurements of one randomly selected gene. Figure 1 displays the gene expression profiles at each sampled point across the 34 replicates.
 - (a) (10') Can you use functional principal component analysis to explore the major modes of variability among these 34 curves?
 - (b) (10') Can you clustering the 34 curves based on their functional principal component scores?

References

Rangel, C., Angus, J., Ghahramani, Z., Lioumi, M., Sotheran, E., Gaiba, A., Wild, D. L., and Falciani, F. (2004). Modeling T-cell activation using gene expression profiling and state-space models. *Bioinformatics*, **20**(9), 1361–1372.

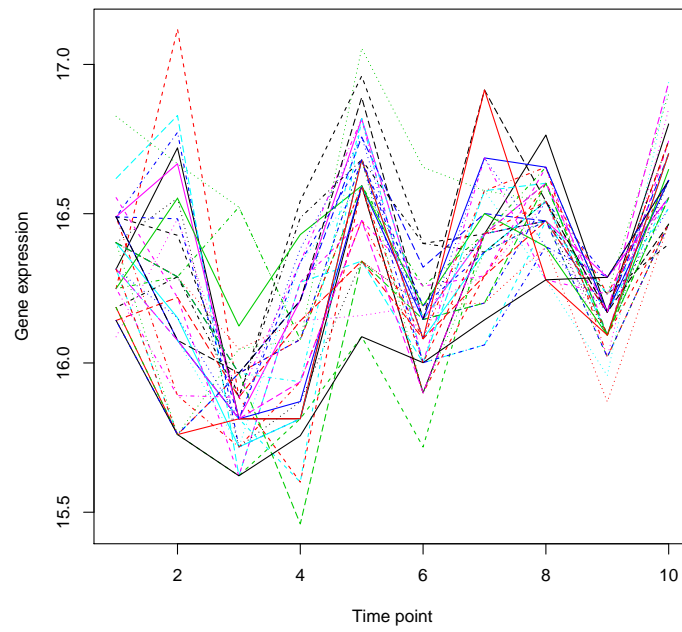


Figure 1: The gene expression of one selected gene.