

Supplementary Information on “Survival Analysis by Penalized Regression and Matrix Factorization”

Yeuntyng Lai, Morihiro Hayashida and Tatsuya Akutsu

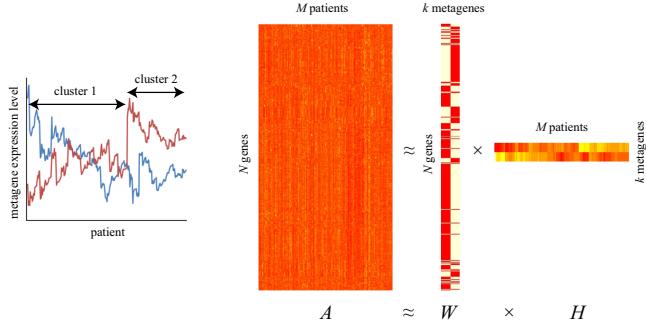


Figure S1: A rank-2 NMF example in microarray data (see also J.P. Brunet, P. Tamayo, T.R. Golub, and J.P. Mesirov, “Metagenes and molecular pattern discovery using matrix factorization,” *PNAS*, vol. 101, pp. 4164–4169, 2004.)

$$c_{12} \begin{pmatrix} 1 & 0 & 0 & 1 & \cdots & 1 & 0 \\ 0 & 1 & & & & & \\ 0 & & 1 & & & & \\ 1 & & & 1 & & & \\ \vdots & & & & \ddots & & \\ 1 & & & & & 1 & 1 \end{pmatrix} \text{ run 1 } \quad \begin{pmatrix} 1 & 1 & 0 & 0 & \cdots & 1 & 0 \\ 1 & 1 & & & & & \\ 0 & & 1 & & & & \\ 0 & & & 1 & & & \\ \vdots & & & & \ddots & & \\ 1 & & & & & 1 & 1 \end{pmatrix} \text{ run 2 } \quad \dots \quad \begin{pmatrix} 1 & 1 & 0 & 1 & \cdots & 1 & 0 \\ 1 & 1 & & 1 & & & \\ 0 & & 1 & & & & \\ 1 & & & 1 & & & \\ \vdots & & & & \ddots & & \\ 1 & & & & & 1 & 1 \end{pmatrix} \text{ run 100 }$$

Figure S2: Connectivity matrix (c_{ij}) over 100 runs.

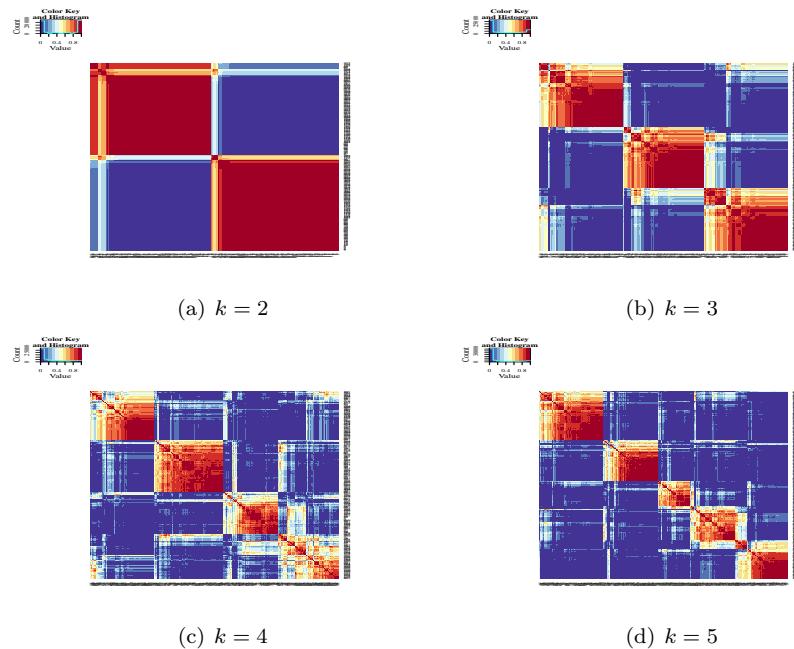


Figure S3: Consensus clustering results of 434 genes without missing values for rank $k = 2, \dots, 5$.

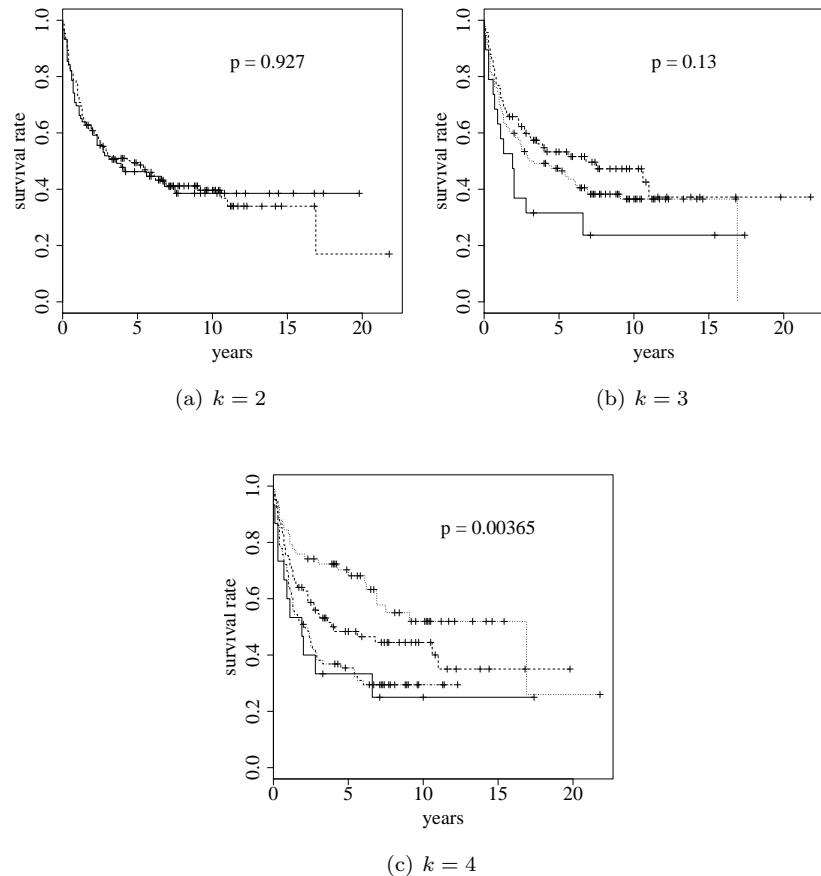


Figure S4: The Kaplan-Meier curve results of three kinds of patient division using 434 genes without missing values corresponding to rank $k = 2, \dots, 4$.