

Advanced Data Analysis (データ解析特論)

Masashi Sugiyama
(Department of Computer Science)
杉山 将 (計算工学専攻)

Contact: W8E-505

sugi@cs.titech.ac.jp

<http://sugiyama-www.cs.titech.ac.jp/~sugi>

Homework

1. Prepare a high dimensional data set and explain the specification of the data.

- Samples should be real-valued **vectors**!
- Better if samples are from your own research area.
- Better if dimensionality is not so small but not too large (say 10 to 100).
- Better if the number of samples is large (say $\gg 100$).
- If your data samples are not vectors (say sequences, images, graphs, texts, etc.), you may use some feature extraction software (developed in your research area) for converting them into vectors.
- You do not have to finalize your data set now; it will be used in the final assignment. Use this opportunity to start searching good data sets.

Homework (cont.)

2. Prepare a computer environment where you can solve **eigenproblems**:
e.g., MATLAB, octave, scilab, R, etc.
- Deadline: **next class** (or e-mail me before the class)

Textbook

- Handouts are provided if necessary.
- Pointers to relevant articles will be provided.
- The following reference may be useful for learning elementally (and advanced) matrix calculations.
 - Matrix Cookbook: <http://matrixcookbook.com/>