

ソフトウェア演習 Ia レポート課題 3 - 1

問題：

ある音声データを特徴量化したデータがテキストファイルとして与えられている。ファイル名は、A01M0007-0007.dat である。データの最初の行には、フレーム数（データの行数）と、各フレームの特徴次元数が書かれている。2 行目以降が実際のデータであり、1 行に 1 フレームのデータが入っている。データはスペースで分割されている。

フレーム数と次元数に応じて動的に 2 次元配列領域を確保し、このデータを読み込んだ上で、このデータファイルに含まれるデータの特徴次元毎の平均からなる平均ベクトルと特徴次元間の分散共分散行列を計算するプログラムを作成しなさい。

ただし、平均ベクトルと分散共分散行列を求める部分は、以下の関数インタフェースで関数化すること。

```
double *get_mean(int ndim, double **data, int ndata)
double **get_cova(int ndim, double **data, int ndata)
```

ここで、ndim はデータの各フレームの次元数（与ファイルデータでは 10）、data

$$\boldsymbol{\mu} = \frac{1}{N} \sum_{n=1}^N \mathbf{x}_n = \begin{bmatrix} \frac{1}{N} \sum_{n=1}^N x_n^1 \\ \frac{1}{N} \sum_{n=1}^N x_n^2 \\ \vdots \\ \frac{1}{N} \sum_{n=1}^N x_n^D \end{bmatrix} \quad D: \text{ベクトルの次元数、} N: \text{データ数}$$
$$\Sigma = \frac{1}{N} \sum_{n=1}^N (\mathbf{x}_n - \boldsymbol{\mu})(\mathbf{x}_n - \boldsymbol{\mu})^t = \begin{bmatrix} \sigma_{11}^2 & \sigma_{12}^2 & \cdots & \sigma_{1D}^2 \\ \sigma_{21}^2 & \sigma_{22}^2 & & \vdots \\ \vdots & & \ddots & \vdots \\ \sigma_{D1}^2 & \cdots & \cdots & \sigma_{DD}^2 \end{bmatrix}$$
$$\sigma_{ij}^2 = \frac{1}{N} \sum_{n=1}^N (x_n^i - \mu^i)(x_n^j - \mu^j)$$

図1 分散共分散行列

はデータが格納されている 2 次元配列、ndata はデータの個数 (フレーム数) である。
両関数とも、必要なメモリ領域は関数内で確保し、確保したメモリのアドレスを関数の
戻り値とする。

レビュー発表担当：

各グループの番号 5 の人が発表しなさい。

提出ファイル名：

XXXXXXreport0301.txt