Numerical Analysis, problem set for final project 2025
Due to 17:00, 5 Jan.
Please submit solutions to email, jmwu@gms.ndhu.edu.tw

1. (15 points) Let X be an $N \times d$ matrix and Y be an $M \times d$ matrix. Write Matlab codes for the following tasks.
   A. Use nested for–loop to calculate a distance matrix D1 whose element D1(n, m) denotes the Euclidean distance between X(n,:) and Y(m,:).
   B. Use no for–loop to calculate another distance matrix D2 whose element D2(n, m) also denotes the Euclidean distance between X(n,:) and Y(m,:)..
   C. For N = 5000, M = 3000 and d= 300, calculate the mean of absolute elements of the difference between D1 and D2, and give the executing time of the two approaches respectively for computing D1 and D2.
2. (70 points) Let $p_m(x) = p(x \mid y_m, A)$ be a multivariate normal pdf,
$$p(x \mid y_m, A) = \frac{1}{(2\pi)^{\frac{d}{2}}\sqrt{|A^{-1}|}} exp(-\frac{(x - y_m)^t A(x - y_m)}{2}),$$
   where $y_m \in R^d$ and A denotes the inverse of a positive–definite matrix. Consider the stochastic generative model in figure 1. Each time a pdf in the model is selected according to probabilities $\pi_1, \ldots, \pi_M$ to generate an instance $x[t]$, where t denotes the time index. Let $\delta[t] = (\delta_1[t], \ldots, \delta_M[t])^T$ be a unitary column vector, where $\delta_m[t] \in \{0,1\}$ and $\sum_{m=1}^{M} \delta_m[t] = 1$. If $\delta_k[t]$ is one, it is implied that at time t the selected one is $p_k(x)$ and $x[t]$ is generated from the pdf $p_k(x)$. Given $X = \{x[t]\}_{t=1}^{N}$, this problem address on fitting the generative model to X. The goal is to retrieve all $\delta[t]$ and model parameters. Let $\pi_m = \frac{1}{M}$ for all m in this example.
   A. Let $X_k = \{x[t] \mid \delta_k[t] = 1 \text{ for all t}\}$. State the mathematical framework for fitting $p_k(x)$ to X_k.

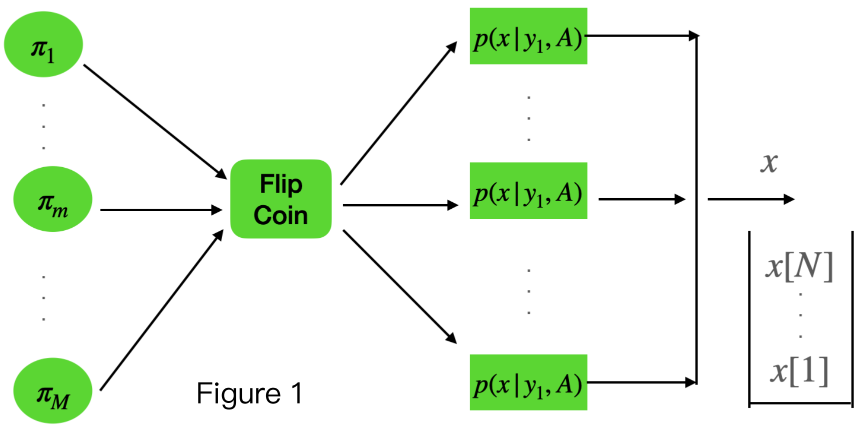

Figure 1

   B. State the mathematical framework for fitting the whole generative model to X. It simply integrates the fitting in problem A.
   C. Derive the rules or equations essential for fitting the whole model.
   D. Draw a flow chart for illustrating your fitting algorithm.

E. Let $y_1 = [z, z]^T, y_2 = [-z, z]^T, y_3 = [z, -z]^T, y_4 = [-z, -z]^T$ and $A = \begin{bmatrix} 0.8 & 0.2 \\ 0.3 & 0.75 \end{bmatrix}$. Let $M = 4$ and $z = 4$. Write Matlab code to simulate the generative model for generating X.

F. Write Matlab codes to implement the flow chart in problem D.

G. Use codes for problem F to fit X that is generated in problem E.

3. (50 points) Let G be a symmetric $N \times N$ matrix that represents connections among $N = 2000$ nodes. Let $G_{ij} = 1$ if nodes i and j are connected and $G_{ij} = 0$ otherwise.

   A. (5 points) $S_i \in \{+1, -1\}$. For graph bisection, let $S_i = 1$ if node i belongs to one set and $S_i = -1$ if node i belongs to the other set. Formulate the constraint on equal size of nodes in two sets.

   B. (5 points) Illustrate the meaning of minimization of $L(S) = -\sum_{i \neq j} S_i G_{ij} S_j$.

   C. (5 points) Derive the objective function, $E(S) = -\sum_{i \neq j} S_i W_{ij} S_j$, where $W_{ij} = G_{ij} - 2\mu$ and $\mu > 0$, for solving the graph bisection problem.

   D. Let S_i be a random variable with possible outcomes belonging $\{+1, -1\}$. Assume $Pr(S_i) \propto exp(-\beta u_i S_i)$, where u_i denotes an auxiliary variable and β denotes the inverse of a positive temperature-like parameter.

      i. (5 points) Express the expectation $\langle S_i \rangle = 1 \times Pr(S_i = 1) + (-1) \times Pr(S_i = -1)$ in terms of u_i and β.

      ii. (5 points) Express the entropy $H(S_i) = -\sum_{S_i = \pm 1} \Pr(S_i) \ln \Pr(S_i)$ in terms of β and u_i.

      iii. (5 points) Express the entropy $H(S_i)$ in terms of $\langle S_i \rangle$.

   E. (10 points) Write the free energy function of mean field approximation in terms of $\langle S_i \rangle$ for all i.

   F. (5 points) Write the mean field equations for solving the graph bisection problem.

   G. (5 points) Draw a flow chart to illustrate the mean field annealing method for solving the graph bisection problem.

4. (40 points) Let $x_i \in R^d$ for $i = 1,\ldots,N$, where d denotes a positive integer. Let $y_k \in R^d$ for $k = 1,\ldots,K$. Both K and N are positive integers and $K < N$. Let $\delta_i = [\delta_{i1}, \ldots, \delta_{iK}]^T$ denote a unitary random vector with $\delta_{ik} \in \{0,1\}$, where $\sum_k \delta_{ik} = 1$ for all i.

   A. (5 points) Let $u_{ik} = -\|x_i - y_k\|$ and e_k denote a unitary vector that consists of K binary elements with only the k^{th} bit one. Assume $\Pr(\delta_i = e_k) \propto \exp(\beta u_{ik})$, where β denotes the inverse of a positive temperature-like parameter. $v_{ik} \equiv \Pr(\delta_i = e_k) = ?$

   B. (5 points) What is the expectation of δ_i ?

   C. (5 points) Derive the entropy of δ_i.

   D. (5 points) Given all $x_i \in R^d$ and v_{ik}, find y_k that minimizes the following criterion, for all k,

   $$E(v, y) = \sum_{i=1}^{N} \sum_{k=1}^{K} v_{ik} \|x_i - y_k\|^2$$

   where v and y respectively collect all v_{ik} and all y_k.

   E. (10 points) Give a flow chart to illustrate the annealed expectation-maximization algorithm for clustering.

   F. (5 points) Describe the halting condition used in the flow chart.

   G. (5 points) Describe the annealing process used in the flow chart.