Percepção

IST – 2004/2005

Guia de Laboratório

Geração de dados: distribuição Gaussiana

Projecto em computador:

Projecto 1: (Baseado no prob 2.24 do livro de reconhecimento de padrões)

Faça uma rotina para simular a geração de vectores de características, produzidos por *c* classes, admitindo que os dados produzidos por cada classe têm distribuição normal multivariada com média e matriz de covariância conhecidas.

Entradas:

Número de vectores de características ($\boldsymbol{N}$ = nº de amostras)
Dimensão de cada vector ($\boldsymbol{d}$)
Número de classes ($\boldsymbol{C}$)
Probabilidade *a priori* de cada classe ($\boldsymbol{P(\omega_i)}$)
Média (vector $\boldsymbol{\mu_i}$) e covariância (matriz $\boldsymbol{R_i}$) dos dados gerados por cada classe

Saídas:

Conjunto de vectores de características
Sequência de classes usadas na geração dos vectores amostra

NOTAS:

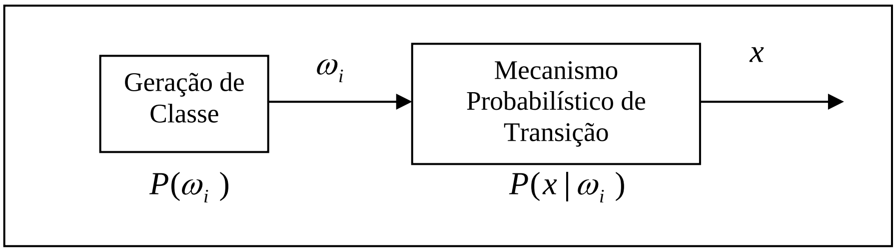
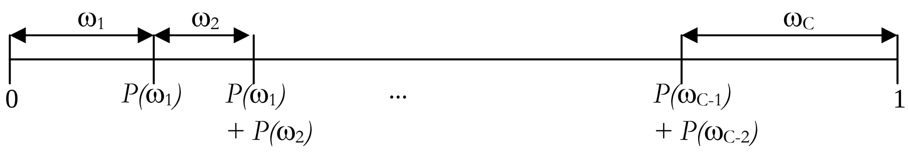


Fig1: Modelo de geração de amostras

A resolução deste trabalho envolve a simulação de dois mecanismos probabilísticos: geração da classe, ω_i, de acordo com uma função de probabilidade conhecida $P(\omega_i)$ e geração de vectores aleatórios com distribuição gaussiana $N(\mu_i, R_i)$, depois de a classe ter sido sorteada

i) *Simulação do mecanismo de geração da classe* ω_i *de acordo com a função de probabilidade conhecida* $P(\omega_i)$: A geração da classe pode ser feita usando uma rotina de geração de números aleatórios com distribuição uniforme, comparando a saída do gerador com a sequência de limiares

$$\lambda_i = \sum_{k=1}^{i} P(\omega_k), \quad i = 1, \ldots, c-1.$$



ii) *Geração de vectores aleatórios de acordo com distribuição Gaussiana* $N(\mu_i, R_i)$ *dado que* ω_i *é conhecido*: A geração de vectores com distribuição normal multivariada, $N(\mu_i, R_i)$, pode ser feita a partir de um gerador de números aleatórios com distribuição gaussiana escalar, $N(0,1)$, com base nos seguintes passos:

- Geração de um vector de variáveis aleatórias independentes, Y, constituído por d saídas do gerador de números aleatórios gaussianos; este vector tem média nula e covariância igual à identidade (N(0,1)).
- Mudança de variável,

$$x = AY + \mu ,$$

em que A é uma matriz quadrada de dimensão d, tal que

$$R = A A^T.$$

$$A = \left(\Lambda^{-\frac{1}{2}} \; \Phi^T \right)^{-1}$$

$$\Lambda = diag(\lambda_1, \lambda_2, \ldots \lambda_d) \quad \text{- valores próprios de } R$$

$$\Phi^T = [e_1, e_2, \ldots e_d] \text{ - vectores próprios de } R$$

Projecto 2: (Prob 2.25 do livro de reconhecimento de padrões)

Use o programa anterior para gerar conjuntos de dados para os seguintes problemas e visualize os resultados.

i) Número de vectores (amostras): 100;
Dimensão de cada vector: $d = 2$;
Número de classes: $c = 2$;
Probabilidades *a priori* : $P(\omega_1) = 0.7, \quad P(\omega_2) = 0.3$;
Distribuição dos vectores de características:

$$N(\mu_i, R_i), \quad i = 1, 2;$$

$$R_i = I, \quad i = 1, 2;$$

$$\mu_1 = [-2.5 \quad 0]^T, \quad \mu_2 = [0 \quad 2.5]^T;$$

ii) Número de vectores (amostras): 1000;
Dimensão de cada vector: $d = 2$;
Número de classes: $c = 9$;
Probabilidades *a priori* : $P(\omega_i) = 1/9, \quad i = 1, \ldots, 9$;
Distribuição dos vectores de características:

$$N(\mu_i, R_i), \quad i = 1, \dots, 9;$$
$$R_i = I, \quad i = 1, \dots, 9;$$
$$\mu_1 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \mu_2 = \begin{bmatrix} 1 \\ 4 \end{bmatrix}, \quad \mu_3 = \begin{bmatrix} 3 \\ 3 \end{bmatrix}, \quad \mu_4 = \begin{bmatrix} 6 \\ -2 \end{bmatrix}, \quad \mu_5 = \begin{bmatrix} 4 \\ 6 \end{bmatrix},$$
$$\mu_6 = \begin{bmatrix} 2 \\ 10 \end{bmatrix}, \quad \mu_7 = \begin{bmatrix} 7 \\ 7 \end{bmatrix}, \quad \mu_8 = \begin{bmatrix} -2 \\ -4 \end{bmatrix}, \quad \mu_9 = \begin{bmatrix} -3 \\ 8 \end{bmatrix}.$$