Классификация на основе оценки апостериорной вероятности

- Как было показано ранее, решение о выборе класса может быть принято на основе оценки апостериорной вероятности, рассчитанной по теореме Байеса, и выбора наибольшей вероятности

$$f(\pi_i \mid X) = \frac{q_i f(X \mid \pi_i)}{\sum_{i=1}^{m} q_i f(X \mid \pi_i)}, \quad i = 1 \ldots m$$

- Априорная вероятность известна из закона многомерного нормального распределения

$$f(X \mid \pi_i) = \frac{1}{(2\pi)^{n/2} \mid \Sigma \mid^{-1}} \exp(-\frac{1}{2}(X - M_i)^T \Sigma^{-1} (X - M_i)) = A \exp((*))$$

Классификация на основе оценки апостериорной вероятности

- Раскроем показатель (*)

$$(*) = -\frac{1}{2}X^T\Sigma^{-1}X + \frac{1}{2}M_i^T\Sigma^{-1}X + \frac{1}{2}M_i^T\Sigma^{-1}M_i - \frac{1}{2}M_i^T\Sigma^{-1}X$$

- Учитывая, что

$$M_i^T\Sigma^{-1}X = X^T\Sigma^{-1}M_i$$

- В результате мы получим следующее выражение

$$f(X|\pi_i) = \exp\left\{-\frac{1}{2}X^T\Sigma^{-1}X\right\}\cdot\exp\{\delta(x)\}$$

$$\text{где}\quad \delta(x) = X^T\Sigma^{-1}M_i - \frac{1}{2}M_i^T\Sigma^{-1}M_i$$

Классификация на основе оценки апостериорной вероятности

- Соответственно, учитывая, что первый сомножитель не зависит от номера класса, так как рассматривается случай равных матриц ковариации, получаем формулу

$$f(\pi_i \mid X) = \frac{q_i e^{\delta_i}}{\sum_{i=1}^{m} q_i e^{\delta_i}}$$

- Решение о классе объекта ищется на основе критерия максимума

$$\max_i \delta_i(x) \quad \text{или} \quad i_{\max} = \arg\max \delta_i$$

Классификация двух нормальных распределений с неравными матрицами ковариации

- Рассмотрим построение статистической решающей функции при условии «$\sum_1$ отлично от $\sum_2$» и заданных математических ожиданиях классов M_1 и M_2
- Построим отношение правдоподобия Λ (X) = f(X | 1) / f(X | 2) , где

$$f(X\,|\,\pi_i)=\frac{1}{(2\pi)^{n/2}\,|\,\Sigma\,|^{-1}}\exp(-\frac{1}{2}(X-M_i)^T\,\Sigma^{-1}(X-M_i))$$

- Рассмотрим отношение правдоподобия

$$\Lambda(X)=\frac{\left|\Sigma_2\right|^{1/2}}{\left|\Sigma_1\right|^{1/2}}\exp\left\{-\frac{1}{2}(X-M_1)^T\,\Sigma_1^{-1}(X-M_1)+\frac{1}{2}(X-M_2)\Sigma_2^{-1}(X-M_2)\right\}\geq K$$

$$K=\frac{C(1\,|\,2)q_2}{C(2\,|\,1)q_1}\quad \textit{известное}\qquad \textit{отношение}$$

Классификация двух нормальных распределений с неравными матрицами распределения

- Логарифмируем отношение правдоподобия

$$\ln\frac{|\Sigma_2|^{1/2}}{|\Sigma_1|^{1/2}}+\frac{1}{2}\Big[(X-M_2)^T\Sigma_2^{-1}(X-M_2)-(X-M_1)^T\Sigma_1^{-1}(X-M_1)\Big]\geq\ln(K)$$

- После приведения к общему виду получим следующую запись

$$\Lambda(X)=\frac{1}{2}X^T\left(\Sigma_2^{-1}-\Sigma_1^{-1}\right)\vec{X}+\vec{X}^T(\Sigma_1^{-1}\vec{M}_1-\Sigma_2^{-1}\vec{M}_2)$$

$$-\frac{1}{2}(\vec{M}_1\Sigma_1^{-1}\vec{M}_1-\vec{M}_2\Sigma_2^{-1}\vec{M}_2)+\ln\frac{|\Sigma_2|^{1/2}}{|\Sigma_1|^{1/2}}\geq\ln K$$

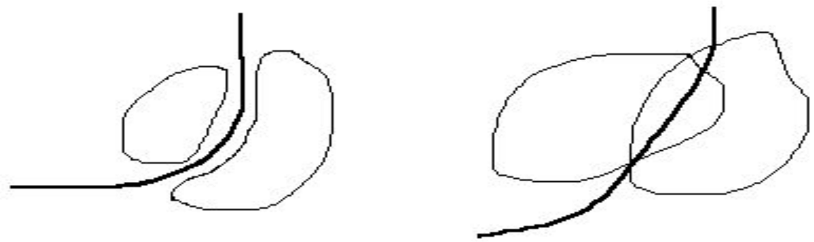
Классификация двух нормальных распределений с неравными матрицами ковариации

- Если выполняется случай $\Sigma_1 = \Sigma_2 = \Sigma$, получаем выражение, изученное ранее

$$X^T \Sigma^{-1}(M_1 - M_2) - \frac{1}{2}(M_1^T \Sigma_1^{-1} M_1 - M_2^T \Sigma_2^{-1} M_2) \geq \ln K$$

- В этом случае мы имеем линейную дискриминацию
- При неравных матрицах ковариации дискриминация будет нелинейной и разделяющая поверхность будет определяться уравнением U(X) = 0
- На следующем рисунке показаны разделяющие поверхности

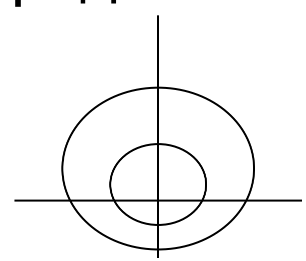
Классификация двух нормальных распределений с неравными матрицами ковариации



- Рассмотрим случай, когда математические ожидания обоих распределений равны нулю : $M_1 = M_2 = 0$. Матрицы ковариации в этом случае определяются следующим образом

$$\Sigma_1 = \begin{pmatrix} \sigma_1^2 & 0 \\ 0 & \sigma_1^2 \end{pmatrix} \qquad \Sigma_2 = \begin{pmatrix} \sigma_2^2 & 0 \\ 0 & \sigma_2^2 \end{pmatrix}$$

- А линии равной плотности представлены на следующем рисунке



Классификация двух нормальных распределений с неравными матрицами ковариации

- В самом частном случае K = 1 $q_1 = q_2$ из чего следует

  ln K = 0 C(2|1) = C(1|2)

$$U(X) = \frac{1}{2} X^T (\Sigma_2^{-1} - \Sigma_1^{-1}) X + \ln \frac{|\Sigma_2|^{1/2}}{|\Sigma_1|^{1/2}} \geq 0$$

$$\Sigma^{-1}{}_1 = \begin{pmatrix} \frac{1}{\sigma_1^2} & 0 \\ 0 & \frac{1}{\sigma_1^2} \end{pmatrix} \qquad \Sigma^{-1}{}_2 = \begin{pmatrix} \frac{1}{\sigma_2^2} & 0 \\ 0 & \frac{1}{\sigma_2^2} \end{pmatrix}$$

$$M_1 = M_2 = \begin{pmatrix} 0 \\ 0 \end{pmatrix} \qquad |\Sigma_1| = \sigma_1^4 \qquad |\Sigma_2| = \sigma_2^4$$

Классификация двух нормальных распределений с неравными матрицами ковариаций

$$\Sigma_2^{-1}-\Sigma_1^{-1}=\begin{pmatrix}\frac{1}{\sigma_2^2}-\frac{1}{\sigma_1^2} & 0\\ 0 & \frac{1}{\sigma_2^2}-\frac{1}{\sigma_1^2}\end{pmatrix} \qquad \ln\frac{\sigma_2^2}{\sigma_1^2}=2\ln\frac{\sigma_2}{\sigma_1}$$

- Теперь подставим все заготовки в общую формулу

$$U(X)=\frac{1}{2}\left(X_1\,,X_2\right)\begin{pmatrix}\frac{1}{\sigma_2^2}-\frac{1}{\sigma_1^2} & 0\\ 0 & \frac{1}{\sigma_2^2}-\frac{1}{\sigma_1^2}\end{pmatrix}\begin{pmatrix}X_1\\ X_2\end{pmatrix}+2\ln\frac{\sigma_2}{\sigma_1}=$$

$$=\frac{1}{2}\left[\frac{\sigma_1^2-\sigma_2^2}{\sigma_1^2\sigma_2^2}(X_1^2+X_2^2)\right]+2\ln\frac{\sigma_2}{\sigma_1}$$

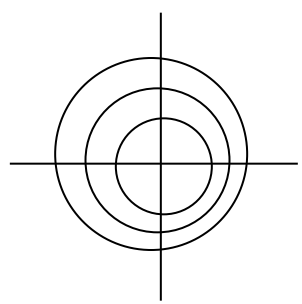
Классификация двух нормальных распределений с неравными матрицами ковариации

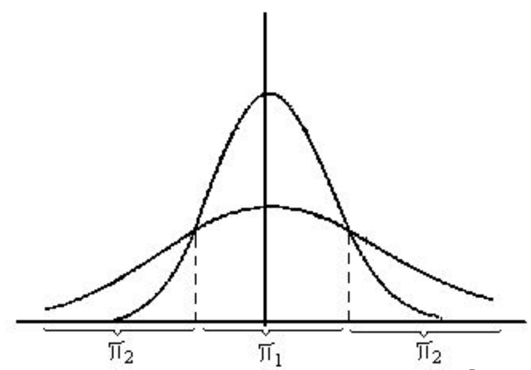
- Решаем уравнение U(x) = 0 и получаем уравнение сферы

$$X_1^2 + X_2^2 = \frac{-4\ln\dfrac{\sigma_2}{\sigma_1}}{\dfrac{1}{\sigma_2^2} - \dfrac{1}{\sigma_1^2}} = R^2 - \textit{так} \quad \textit{обозначим} \quad \textit{радиус полученной сферы}$$

- Таким образом, разделяющая поверхность выглядит очень просто

$$R = 2\sqrt{\frac{\ln\dfrac{\sigma_2}{\sigma_1}}{\dfrac{1}{\sigma_1^2} - \dfrac{1}{\sigma_2^2}}}$$





На рисунке справа показан случай в одномерной области

До Следующей Лекции!