1. Частотный подход к построению статистических моделей. Требования к функции оценивания.

Частотная (классическая или ортодоксальная) статистика основана на концепции выборочного распределения. Это - распределение, что у оценщика δ есть, когда относится несколько набор данных $\mathcal{D}1, \mathcal{D}2 \cdots \mathcal{D}K$ выбранный от истинных, но неизвестных распределений.
Оценка параметра - $\theta\,\hat{} = \delta\,\mathcal{D}$.

Выбрав оценщик, мы определяем его ожидаемую потерю или риск как

$$R(\theta^*, \delta) \triangleq \mathbb{E}_{p(\tilde{\mathcal{D}}|\theta^*)}\left[L\left(\theta^*, \delta\left(\tilde{\mathcal{D}}\right)\right)\right] = \int L\left(\theta^*, \delta\left(\tilde{\mathcal{D}}\right)\right) p\left(\tilde{\mathcal{D}}|\theta^*\right) d\,\tilde{\mathcal{D}},$$

где $\mathcal{D}$ - данные, выбранные "природные" распределения, которые представлено параметром $\theta *$.

2. Квадратический дискриминантный анализ. Порядок получения формулы оценки вероятности p(x|y=c, \theta)

Квадратичный дискриминантный анализ (QDA) является нелинейным обобщением метода LDA, который не использует предположения об однородности ковариационной матрицы.
В качестве решающего правила применяется квадратичная функция

$$d_k = -0.5(\mathbf{x} - \boldsymbol{\mu}_k)^T \mathbf{C}_k^{-1}(\mathbf{x} - \boldsymbol{\mu}_k) - 0.5 \ln |\mathbf{C}_k| + \ln \pi_k,$$

где
$|\mathbf{C}_k|$ - детерминант ковариационной матрицы k-го класса,
$\mathbf{C}_k^{-1}$ - ее обратная матрица;
π_k - априорная вероятность наблюдения объектов k-го класса.
Экзаменуемый объект так же относится к классу с максимальным значением d_k

Gaussian Discriminant Analysis (2/3)

By plugging in the definition of Gaussian density the posteriors over the class labels, we obtain **quadratic discriminant analysis**:

$$\boxed{p(\mathbf{x}|y=c,\boldsymbol{\theta}) = \frac{\pi_c|2\pi\boldsymbol{\Sigma}_c|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_c)^T\boldsymbol{\Sigma}_c^{-1}(\mathbf{x}-\boldsymbol{\mu}_c)\right]}{\sum_{c'}\pi_{c'}|2\pi\boldsymbol{\Sigma}_{c'}|^{-\frac{1}{2}}\exp\left[-\frac{1}{2}(\mathbf{x}-\boldsymbol{\mu}_{c'})^T\boldsymbol{\Sigma}_{c'}^{-1}(\mathbf{x}-\boldsymbol{\mu}_{c'})\right]}}$$

Consider a special case in which the covariance matrices are tied or shared across classes ($\boldsymbol{\Sigma}_c = \boldsymbol{\Sigma}$):

$$\boxed{p(\mathbf{x}|y=c,\boldsymbol{\theta}) \propto \exp\left[\boldsymbol{\mu}_c^T\boldsymbol{\Sigma}^{-1}\mathbf{x} - \frac{1}{2}\boldsymbol{\mu}_c^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_c + \log[\pi_c]\right]\exp\left[-\frac{1}{2}\mathbf{x}^T\boldsymbol{\Sigma}^{-1}\mathbf{x}\right]}$$

Let us define:

$$\boxed{\gamma_c = -\frac{1}{2}\boldsymbol{\mu}_c^T\boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_c + \log[\pi_c]} \qquad \boxed{\boldsymbol{\beta}_c = \boldsymbol{\Sigma}^{-1}\boldsymbol{\mu}_c}$$

Gaussian Discriminant Analysis (3/3)

Then we can write:

$$\boxed{p(\mathbf{x}|y=c,\boldsymbol{\theta}) = \frac{\exp[\boldsymbol{\beta}_c^T\mathbf{x}+\gamma_c]}{\sum_{c'}\exp[\boldsymbol{\beta}_{c'}^T\mathbf{x}+\gamma_{c'}]} = \mathcal{S}(\boldsymbol{\eta})_c}$$

where $\boldsymbol{\eta} = \left[\boldsymbol{\beta}_1^T\mathbf{x}+\gamma_1,\cdots,\boldsymbol{\beta}_C^T\,\mathbf{x}+\gamma_C\right]$, and $\mathcal{S}$ is softmax function defined as:

$$\boxed{\mathcal{S}(\boldsymbol{\eta})_c = \frac{e^{\eta_c}}{\sum_{c'=1}^{C}e^{\eta_{c'}}}}$$

If we take logs, we end up with linear function of $\mathbf{x}$ (because $\mathbf{x}^T\boldsymbol{\Sigma}^{-1}\mathbf{x}$) cancels from numerator/denominator). Thus the decision boundary between any two classes well be a straight line. Hence this technique is called **linear discriminant analysis (LDA).**

3. Определенить параметры модели GMM по заданной эмпирической выборке:

X = {(1,2); (1,4); (-4, 0); (0, 10); (3, 5); (-7, 12)}

Mixtures of Gaussians (***Gaussian Mixture Model***):

$$\boxed{p(\mathbf{x}_i|\boldsymbol{\theta}) = \sum_{k=1}^{K}\pi_k\mathcal{N}(\mathbf{x}_i|\boldsymbol{\mu}_k,\boldsymbol{\Sigma}_k)}$$