№35.

$L(y', y) = (y' - y)^2$

Д-ть: если $f^*(x) = \underset{c}{\operatorname{argmin}} E[(Y-c)^2 [X = x]$,

то $f^*(x) = E[Y | X = x]$. Чему равен средний риск $R(f^*)$?

$\rceil\ y \in Y = \mathbb{R}$, $L(f(x), y) = (f(x) - y)^2$
$(X = X)$

$\Rightarrow$ Ф-ал среднего риска имеет следующий вид:

$$R(f) = \int_{X \times Y} L(f(x), y) \cdot p(x, y)\, dx\, dy = \int_X \boxed{\int_Y L(f(x), y) \cdot p(y | x)\, dy}\, p(x)\, dx =$$

$$= \int_X \underset{\overset{\shortparallel}{0}}{p(x)} \cdot \underline{\left[\int_Y (f(x) - y)^2 \cdot p(y | x)\, dy \right]} dx$$

$\hookrightarrow$ необходимо минимизировать внутрений интеграл, т.к тогда минимизируется и $R(f)$

$\hookrightarrow$ (т.к. $p(x) \geq 0$). Минимизировать можно поточечно:

Для фиксирован. x найдем значение $c = f(x_0)$, при котором

$$\underbrace{\int_Y (c - y)^2\, p(y | x_0)\, dy}_{\hat{=}\, E_y[(c-y)^2 | x_0]} \to \min_c$$

$$\frac{\partial}{\partial c} E_y[(c-y)^2 | x_0] = 0$$

$$\frac{\partial}{\partial c} \int_Y (c-y)^2\, p(y | x_0)\, dy = 2 \int_Y (c-y)\, p(y | x_0)\, dy = 2\left[\int_Y c\, p(y | x_0)\, dy - \right.$$

$$\underbrace{-\int_{\mathcal{Y}} y\, p(y|x_0)\,dy}_{\triangleq E_y[y|x_0]}] = 2\Big[c\underbrace{\int_{\mathcal{Y}} p(y|x_0)\,dy}_{=1} - E_y[y|x_0]\Big] =$$

(по св-ву условного распределения)

$$= 2(c - E_y[y|x_0]) = 0$$

$$\int_{\mathcal{Y}} p(y|x_0)\,dy = P(y \in \mathcal{Y}|x_0) = 1$$ – т.к при $x = x_0$, y примет одно из своих возможных значений $\mathcal{Y}$ $\Rightarrow$ это достоверное событие

$$2(c - E_y[y|x_0]) = 0 \Rightarrow \boxed{c = E_y[y|x_0]}$$

Т.к мы фиксировали произвольный x_0, то полученный рез-т можно обобщить $\forall x \in \mathcal{X}$ присем тогда $c \Rightarrow c(x) = f^*(x)$, т.е c – не фиксировано, а зависит от x.

$$c = E_y[y|x_0] \Rightarrow c(x) = f^*(x) = E_y[y|x] \;\#$$

$\Rightarrow$ средний риск $R(f^*) = \int_{\mathcal{X}} p(x)\Big[\int_{\mathcal{Y}} (f^*(x) - y)^2 \cdot p(y|x)\,dy\Big]dx$

$$\int_{\mathcal{Y}} (f^*(x) - y)^2 \cdot p(y|x)\,dy = \int_{\mathcal{Y}} (E_y[y|x] - y)^2 \cdot p(y|x)\,dy =$$
$$= \int_{\mathcal{Y}} (y - E_y[y|x])^2 \cdot p(y|x)\,dy = E_y[(y - E_y[y|x])^2 | x]$$

$$E_y[(y - E_y[y|x])^2 | x] \triangleq D_y[y|x]$$

$$R(f^*) = \int_{\mathcal{X}} p(x)\, D_y[y|x]\,dx \triangleq E_x[D_y[y|x]]$$

Ответ: $f^*(x) = E_y[y|x]$; $R(f^*) = E_x[D_y[y|x]]$

№36.

$\mathcal{L}(y', y) = |y' - y|$

Доказать: $f^*(x) = \text{median}(Y / X = x) \Leftrightarrow R(f^*) \to \min$

$\square$: $Y \in \mathbb{Y} = \mathbb{R}$; $\mathcal{L}(f(x), y) = |f(x) - y| \Rightarrow$ ф-ал среднего риска имеет вид:
$(X = \mathbb{X}$

$$R(f) = \int_{\mathbb{X}\times\mathbb{Y}} \mathcal{L}(f(x),y)\cdot p(x,y)\,dx\,dy = \int_{\mathbb{X}} p(x) \boxed{\int_{\mathbb{Y}} \mathcal{L}(f(x),y)\cdot p(y|x)\,dy}\,dx$$

· Минимизируем внутренний интеграл, потому что тогда минимизируется и $R(f)$ ($\because p(x) \geq 0$). Минимизировать можно поточечно. Для фиксированного x_0 найдем значение $c = f(x_0)$, при котором $\boxed{\int_{\mathbb{Y}} \mathcal{L}(c,y)\cdot p(y|x_0)\,dy} \to \min$

$= E_y[\mathcal{L}(c,y)|x_0]$

$$\frac{\partial}{\partial c} E_y[\mathcal{L}(c,y)|x_0] = 0$$

$$\frac{\partial}{\partial c}\int_{\mathbb{Y}} \mathcal{L}(c,y)\,p(y|x_0)\,dy = \frac{\partial}{\partial c}\int_{\mathbb{Y}}|c-y|\cdot p(y|x_0)\,dy = \frac{\partial}{\partial c}\int_{\mathbb{Y}\setminus\{c\}}|c-y|\,p(y|x_0)\,dy =$$

$$= \int_{\mathbb{Y}\setminus\{c\}} \text{sign}(c-y)\cdot p(y|x_0)\,dy = \int_{\{c>y\}} \overbrace{\text{sign}(c-y)}^{1} p(y|x_0)\,dy +$$

$$+ \int_{\{c<y\}} \overbrace{\text{sign}(c-y)}^{=-1} p(y|x_0)\,dy = \int_{\{c>y\}} p(y|x_0)\,dy - \int_{\{c<y\}} p(y|x_0)\,dy =$$

$$= \underbrace{\boxed{\int P(\{c>y\}|x_0)}}_{= P(\{c>y\}|x_0)} - \underbrace{\boxed{P(\{c<y\}|x_0)}}_{= P(\{c<y\}|x_0)} = 0$$

$$\Rightarrow P(\{c>y\}|x_0) = P(\{c<y\}|x_0)$$

Будем считать, что $p(y|x)$ непрерывно. Тогда,

$P(\{c=y\}|x_0) = 0 \Rightarrow$

$$\begin{cases} P(\{c>y\}|x_0) = \frac{1}{2} \\ P(\{c<y\}|x_0) = \frac{1}{2} \end{cases} \Rightarrow c = \mathrm{median}(Y|X=x_0)$$

Т.к фиксирован произвольный x_0, то рез-т можно обобщить для $\forall x \in X$, причем тогда:

$c \to c(x) = f^*(x)$, т.е c зависит от x:

$c = \mathrm{median}(Y|X=x_0) \to c(x) = f^*(x) = \mathrm{median}(Y|X=x)$ #

№37.

$\mathcal{L}(y', y) = ?: R(f) \to \min\limits_f$ при $f^*(x) = \mathrm{mode}(Y|X=x)$

Из предыдущих задач: $R(f) \to \min\limits_f \Leftrightarrow E_y[\mathcal{L}(c,y)|x_0] \to \min\limits_c$

т.е $\boxed{\int_{\mathbb{R}} \mathcal{L}(c,y) \cdot p(y|x_0)\,dy \to \min\limits_c}$ (1)

$f^*(x) = \mathrm{mode}(Y|X=x) \Rightarrow p(c|x) \to \max\limits_c$

Т.е при фиксированном x_0 и $f^*(x_0) = c^*$:

$c^* = \mathrm{mode}(y|x_0) \Rightarrow p(c^*|x_0) = \max\limits_y p(y|x_0)$, что можно записать так: $c^* = \mathrm{mode}(y|x_0) \Rightarrow \boxed{-p(c^*|x_0) = \min\limits_c (-p(c|x_0))}$ (2)

Учитывая (1) и (2): $\boxed{\int_{\mathbb{R}} \mathcal{L}(c,y) \cdot p(y|x_0)\,dy = -p(c|x_0)}$

$\exists\ \hat{\mathcal{L}}(c,y) = -\mathcal{L}(c,y) \Rightarrow \int_{\mathbb{R}} \hat{\mathcal{L}}(c,y) \cdot p(y|x_0) = p(c|x_0)$

$\exists$ ф-я, удовлетв. усл: $\hat{\mathcal{L}}(c,y) = \delta(c-y) = \delta(y-c)$ дельта-ф-я Дирака

Она обладает св-ом: $\boxed{\int_{\mathbb{R}} f(x)\,\delta(x-x_0)\,dx = f(x_0)}$

$\Rightarrow \mathcal{L}(c,y) = -\delta(y-c)$, но: $\mathcal{L}(c,y) \geq 0$, а $\delta(y-c) \geq 0$

$\Rightarrow \exists\, \underbrace{c_0 > 0}_{\to \text{ввели}}$.

$\Rightarrow \boxed{\mathcal{L}(c,y) = c_0 - \delta(y-c)} \Rightarrow E_y[\mathcal{L}(c,y)|x_0] = \int\limits_{\mathbb{R}} \mathcal{L}(c,y) \cdot p(y|x_0)\,dy =$

$= \int\limits_{\mathbb{R}} c_0 p(y|x_0)\,dy - \underbrace{\int\limits_{\mathbb{R}} \delta(y-c) p(y|x_0)\,dy}_{= p(c|x_0)} =$

$= c_0 \underbrace{\int\limits_{\mathbb{R}} p(y|x_0)\,dy}_{=1} - p(c|x_0) = c_0 - p(c|x_0) \to \min\limits_{c} \Rightarrow p(c|x_0) \to \max$

$\Rightarrow c = f^*(x_0) = \mathrm{mode}(y|x_0)$ #

Ответ: $\boxed{\mathcal{L}(y', y) = c_0 - \delta(y - y');\ c_0 > 0}$

№4.1

Д-ть что можно положить $v_0 = \overline{x}$. Описать мн-во всех возможных v_0 доставляющих минимум суммы квадратов расстояний до искомого многообразия.

Решение:

$W_0 = v_0 \in \mathbb{R}^d$

$\sum\limits_{i=1}^{N} \mathrm{dist}^2(x^{(i)}, W_0) = \sum\limits_{i=1}^{N} \|x^{(i)} - v_0\|^2 \to \min\limits_{v_0 \in \mathbb{R}^d}$

$\frac{d}{dv_0}\left(\sum\limits_{i=1}^{N} \|x^{(i)} - v_0\|^2\right) = \sum\limits_{i=1}^{N} \frac{d(\|x^{(i)} - v_0\|^2)}{dv_0} = \sum\limits_{i=1}^{N} \frac{d(\|x^{(i)} - v_0\|^2)}{d(x^{(i)} - v_0)} \cdot \frac{d(x^{(i)} - v_0)}{dv_0} =$

$= \sum\limits_{i=1}^{N} 2(x^{(i)} - v_0) \cdot (-1) = 0 \;\searrow\; \sum\limits_{i=1}^{N} (x^{(i)} - v_0) = 0$

$\sum\limits_{i=1}^{N} x^{(i)} - \sum\limits_{i=1}^{N} v_0 = 0 \Rightarrow \sum\limits_{i=1}^{N} x^{(i)} - N v_0 = 0$

Ответ: $v_0 = \frac{\sum\limits_{i=1}^{N} x^{(i)}}{N}$

№17.

X_1	4	0	-1	3	4
X_2	2	-3	-2	1	2
X_3	3	2	2	1	-3

Найти гл. направления и дисперсию по гл. компонентам
Изобразить точки и гл. направл.

Решение:

$$X = \begin{pmatrix} 4 & 2 & 3 \\ 0 & -3 & 2 \\ -1 & -2 & 2 \\ 3 & 1 & 1 \\ 4 & 2 & -3 \end{pmatrix} \to \bar{x} = (2, 0, 1)$$

$$X_c = \begin{pmatrix} 2 & 2 & 2 \\ -2 & -3 & 1 \\ -3 & -2 & 1 \\ 1 & 1 & 0 \\ 2 & 2 & -4 \end{pmatrix}$$ — центрировани данные

$$C = X_c^T X_c : \begin{pmatrix} 2 & -2 & -3 & 1 & 2 \\ 2 & -3 & -2 & 1 & 2 \\ 2 & 1 & 1 & 0 & -4 \end{pmatrix} \begin{pmatrix} 2 & 2 & 2 \\ -2 & -3 & 1 \\ -3 & -2 & 1 \\ 1 & 1 & 0 \\ 2 & 2 & -4 \end{pmatrix} = \begin{pmatrix} 22 & 21 & -9 \\ 21 & 22 & -9 \\ -9 & -9 & 22 \end{pmatrix}$$

$$\frac{1}{N-1} C = \begin{pmatrix} 5.5 & 5.25 & -2.25 \\ 5.25 & 5.5 & -2.25 \\ -2.25 & -2.25 & 5.5 \end{pmatrix}$$ — выборочная матрица ковариации

находим с.з и с.в:

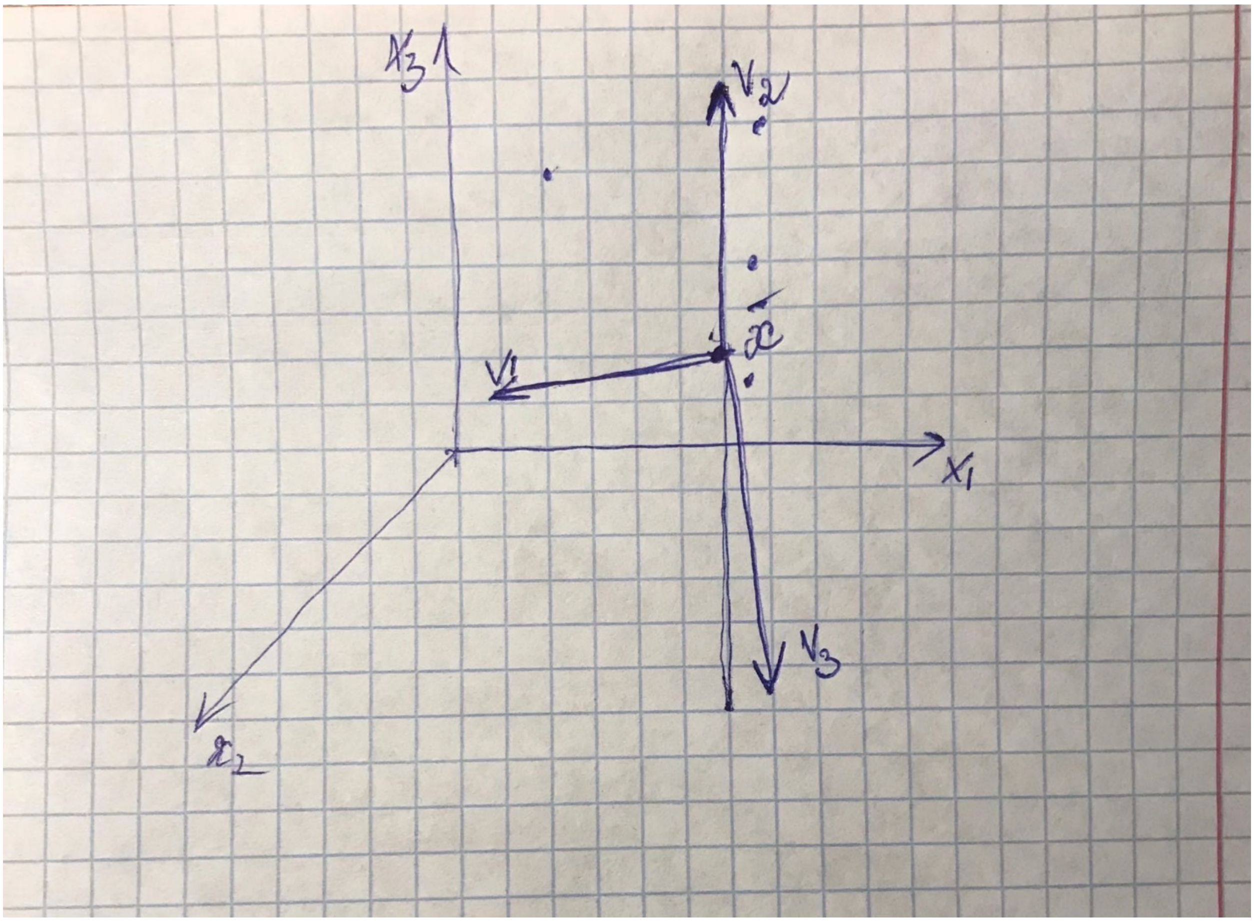
$$\begin{pmatrix} 22-\lambda & 21 & -9 \\ 21 & 22-\lambda & -9 \\ -9 & -9 & 22-\lambda \end{pmatrix} \to \lambda_1 = 1 \quad \lambda_2 = 16 \quad \lambda_3 = 49$$

$$V_1 = \frac{\sqrt{2}}{2} \begin{pmatrix} -1 \\ 1 \\ 0 \end{pmatrix} \quad V_2 = \frac{\sqrt{11}}{11} \begin{pmatrix} 1 \\ 1 \\ 3 \end{pmatrix} \quad V_3 = \frac{\sqrt{22}}{22} \begin{pmatrix} 3 \\ 3 \\ -2 \end{pmatrix}$$

это главные компоненты

Дисперсии по глав. компон:

$$\frac{1}{N-1} \lambda_1 = 0.25 \quad \frac{1}{N-1} \lambda_2 = 4 \quad \frac{1}{N-1} \lambda_3 = 12.25$$



x3
v2
x
v1
x1
v3
x2