№1

Курасбедиани Г.Г.
3824М1ПМки

Пусть $f: \mathbb{R}^n \to \mathbb{R}^n$, $x \in \mathbb{R}^n$. Доказать:

1) если $a \in \mathbb{R}^n$, $x \in \mathbb{R}^n$, то $\frac{\partial(a^T x)}{\partial x} = a$

2) если $A \in \mathbb{R}^{m \times n}$, $x \in \mathbb{R}^n$, то $\frac{\partial(Ax)}{\partial x} = A$

3) если $A \in \mathbb{R}^{n \times n}$, $x \in \mathbb{R}^n$, то $\frac{\partial(x^T A x)}{\partial x} = (A + A^T)x$;
если $A^T = A$, то $\frac{\partial(x^T A x)}{\partial x} = 2Ax$

4) если $x \in \mathbb{R}^n$, то $\frac{\partial \|x\|^2}{\partial x} = 2x$

5) если f – скал. ф-я и под $f(x)$ понимается покомпонентное применение f к $x \in \mathbb{R}^n$, то $\frac{\partial f(x)}{\partial x} = diag(f'(x))$

6) если $h: \mathbb{R}^n \to \mathbb{R}^m$, $g: \mathbb{R}^m \to \mathbb{R}^p$, $x \in \mathbb{R}^n$, то $\frac{\partial g(h(x))}{\partial x} = \frac{\partial g(h(x))}{\partial h} \frac{\partial h(x)}{\partial x}$

1) $a^T x = \sum\limits_{i=1}^n a_i x_i$

$$\frac{\partial(a^T x)}{\partial x} = \frac{\partial}{\partial x} \sum_{i=1}^n a_i x_i = \underset{j=\overline{1,n}}{\frac{\partial}{\partial x_j}} \sum_{i=1}^n a_i x_i = \left[\begin{matrix} a_i, & i = j \\ 0, & i \neq j \end{matrix}\right] = (a_1, a_2, \dots, a_n) = a$$

2) $$Ax = \begin{pmatrix} \sum\limits_{i=1}^n a_{1i} x_i \\ \vdots \\ \sum\limits_{i=1}^n a_{mi} x_i \end{pmatrix}$$

Для каждой строки аналогично п.1 $\Rightarrow$ получаем, что $\frac{\partial(Ax)}{\partial x} = A$

3) $x^T A x = (Ax, x)$

$[D_{x_0} f](h)$ – дифференциал ф-и f в т. x_0, h – малое приращение

$$[D_{x_0}(Ax, x)](h) = ([D_{x_0} Ax](h), x_0) + (Ax_0, [D_{x_0} x](h)) = (Ah, x_0) + (Ax_0, h)$$

$$(Ah, x_0) + (Ax_0, h) = (A^T x_0, h) + (Ax_0, h) = ((A^T + A)x_0, h)$$

Получаем $\partial \frac{x^T A x}{\partial x} = (A^T + A)x$; если $A^T = A$, то $2Ax$

Курастбдиали Б.И
3824М1ПМкч

4) $\|x\|^2 = x^T x = (x, x)$

$$[D_{x_0}(x,x)](h) = \left([D_{x_0} x](h), x\right) + \left(x, [D_{x_0} x](h)\right) = (h, x) + (x, h)$$

$$(h, x) + (x, h) = (2x, h)$$

Получаем $\dfrac{\partial \|x\|^2}{\partial x} = 2x$

5) $$\frac{\partial g(x)}{\partial x} = \left(\frac{\partial g(x_i)}{\partial x_j}\right) = \begin{bmatrix} g'(x_i), & i = j \\ 0, & i \neq j \end{bmatrix}$$

$$\frac{\partial g(x)}{\partial x} = \begin{pmatrix} g'(x_1) & & \\ & \ddots & \\ & & g'(x_n) \end{pmatrix}$$

6) $$\frac{\partial g(h(x))}{\partial x} = \left(\frac{\partial g_i(h(x))}{\partial x}\right) = \left(\sum_{j=1}^{m} \frac{\partial g_i}{\partial h_j} \frac{\partial h_j}{\partial x}\right)$$

$\rceil$ $f(x) = g(h(x))$

$$f(x_0 + h) - f(x_0) = g(h(x_0 + \Delta x)) - g(h(x_0)) \approx$$

$$\approx [D_{h(x_0)} g](h(x_0 + \Delta x) - h(x_0)) \approx [D_{h(x_0)} g]([D_{x_0} h](\Delta x))$$

Получаем $\dfrac{\partial g(h(x))}{\partial x} = \dfrac{\partial g(h(x))}{\partial h} \dfrac{\partial h(x)}{\partial x}$

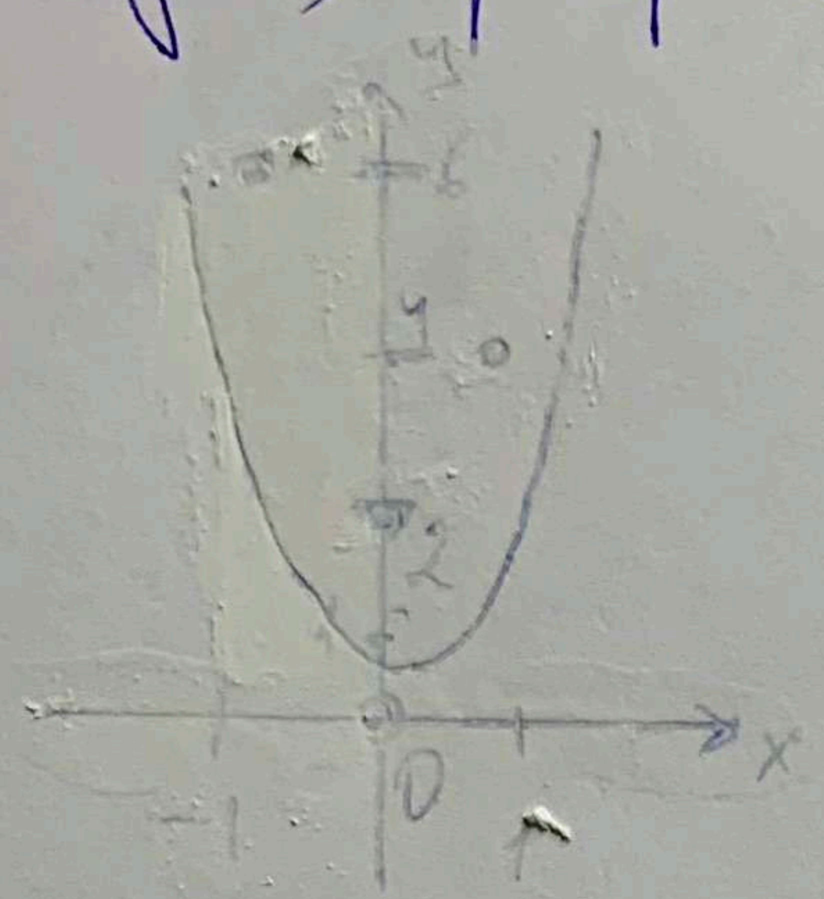
Курасбелиани Т.И
3824М1ПМки

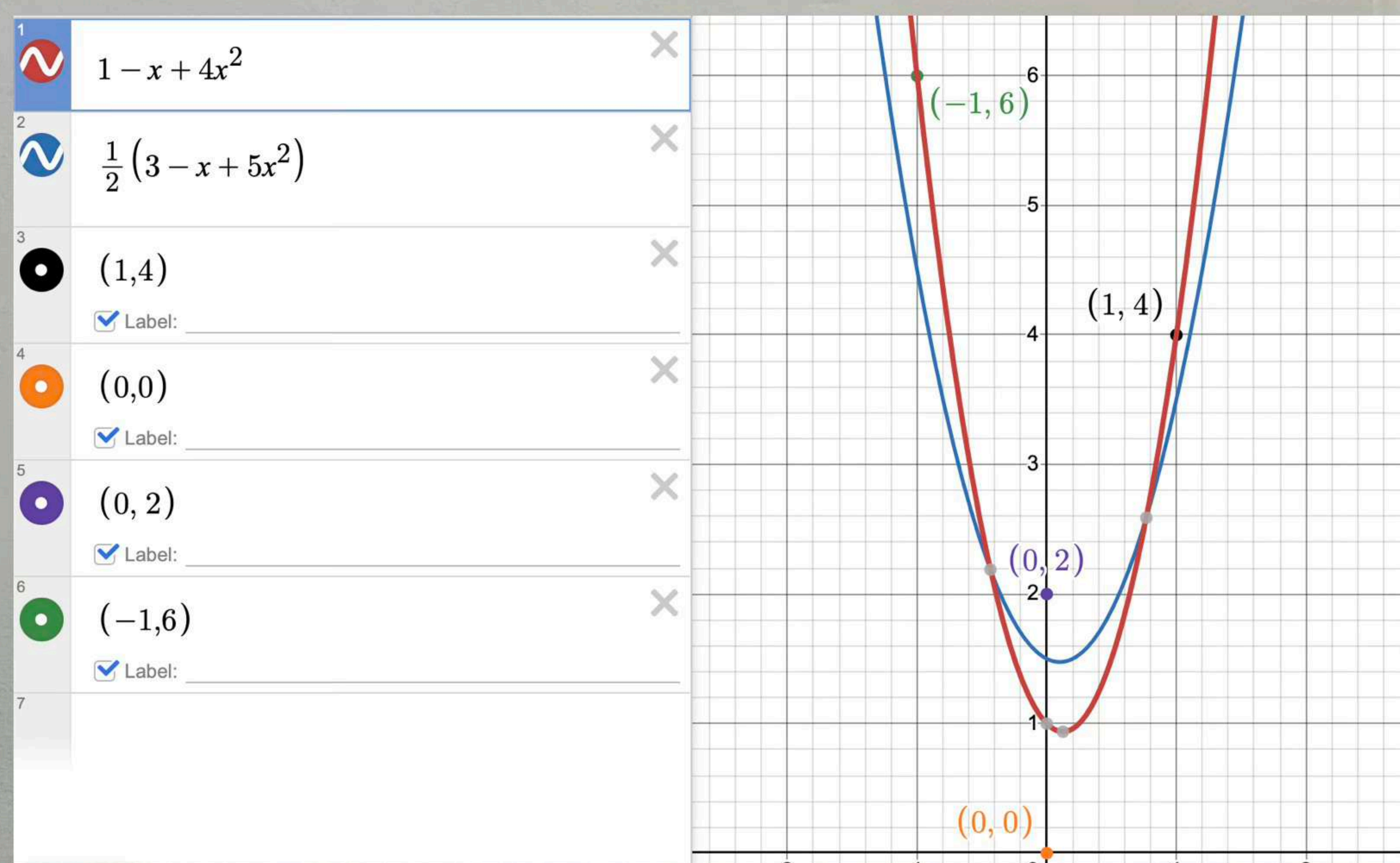
№3

x	1	1	0	0	−1
y	4	4	0	2	6

1) Изобразить точки
2) МНК построить модель вида $f(x)=\sum_{i=0}^{2}\beta_i x^i$. Построить график
3) Ridge; график ($\lambda=1$)

1)





2) $X=\begin{pmatrix}1&1&1\\1&1&1\\1&0&0\\1&0&0\\1&-1&1\end{pmatrix}$ $y=\begin{pmatrix}4\\4\\0\\2\\6\end{pmatrix}$

$X^TX=\begin{pmatrix}5&1&3\\1&3&1\\3&1&3\end{pmatrix}$ $X^Ty=\begin{pmatrix}16\\2\\14\end{pmatrix}$

$\beta=(X^TX)^{-1}X^Ty=\begin{pmatrix}1\\-1\\4\end{pmatrix}$

$f(x)=1-x+4x^2$

3) $\beta^{ridge}=(X^TX+\lambda I)^{-1}X^Ty=\begin{pmatrix}3/2\\-1/2\\5/2\end{pmatrix}$

$f(x)=\frac{3}{2}-\frac{1}{2}x+\frac{5}{2}x^2$

Курасбедиани Г.И.
3824 М1ПМки

№4

* задачу восстановления регрессии, в которой $y \sim N(X\beta, \sigma^2 I)$, а β имеет априорное распределение $N(0, \tau I)$

1) Найти апостериорное распределение для β.

$$P(\beta/y) \sim P(\beta) P(y/\beta)$$

$$P(\beta/y) = \frac{1}{\sqrt{2\tau\pi}} e^{-\frac{1}{2\tau}\|\beta\|^2} \cdot \frac{1}{\sigma^2\sqrt{2\pi}} e^{-\frac{1}{2\sigma^2}\|y - X\beta\|^2}$$

$$P(\beta/y) = \frac{1}{2\pi\sigma\sqrt{\tau}} e^{-\frac{1}{2}\left(\frac{1}{\tau}\|\beta\|^2 + \frac{1}{\sigma^2}\|y - X\beta\|^2\right)}$$

2) Д-ть, что β^{ridge} есть его мат. ожидание:

$$E = \arg\max_{\beta} e^{-\frac{1}{2}\left(\frac{1}{\tau}\|\beta\|^2 + \frac{1}{\sigma^2}\|y - X\beta\|^2\right)} =$$

$$= \arg\min_{\beta} \left(\|y - X\beta\|^2 + \frac{\sigma^2}{\tau}\|\beta\|^2\right)$$

3) связь: $\lambda = \frac{\sigma^2}{\tau}$

Курасбедиани Г.К
3824М1ПМкн

№5

Показать, что процедура гребневой регрессии экв. МНК, примен. к расшир. данным

$$\beta = (X^T X)^{-1} X^T y$$

$$\beta^{ridge} = (X^T X + \lambda I)^{-1} X^T y$$

$$\hat{X} = \begin{pmatrix} X \\ \sqrt{\lambda} I \end{pmatrix} \in \mathbb{R}^{n+d \times n} \qquad \hat{y} = \begin{pmatrix} y \\ 0 \end{pmatrix} \in \mathbb{R}^{n+d}$$

$$\hat{X}^T \hat{X} = \begin{pmatrix} X^T & \sqrt{\lambda} I \end{pmatrix} \begin{pmatrix} X \\ \sqrt{\lambda} I \end{pmatrix} = X^T X + \lambda I$$

$$\hat{X}^T \hat{y} = \begin{pmatrix} X^T & \sqrt{\lambda} I \end{pmatrix} \begin{pmatrix} y \\ 0 \end{pmatrix} = X^T y$$

$$\hat{\beta} = (\hat{X}^T \hat{X})^{-1} \hat{X}^T \hat{y} = (X^T X + \lambda I)^{-1} X^T y = \beta^{ridge}$$