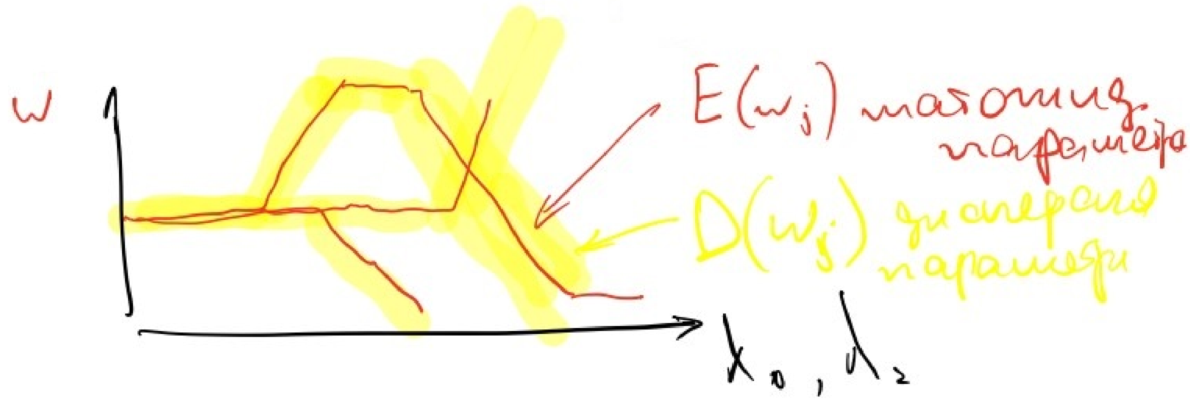
Решается задача регрессии. На подвыборках, равномощных выборке, оцениваются параметры. По K подвыборкам получены матожидание и дисперсия каждого параметра модели. Требуется построить график (λ, ω_j), где λ – коэффициент регуляризации функции ошибки.

1) $S = ||f - y||_2^2 + \lambda_2 ||w||_2^2$
2) $S = ||f - y||_2^2 + \lambda_1 ||w||_1$
3) $S = ||f - y||_2^2 + \lambda_2 ||w||_2^2 + \lambda ||w||_1$



Пусть $F = f(\omega, X) = X_{m \cdot n} w_{n \cdot 1}$ - модель
$(X, y_{m \cdot 1})$ - выборка

1 Решение:

1.1 I

$$S = ||Xw - Y||_2^2 + \lambda_2 ||w||_2^2 \tag{1}$$

$$S = (Xw - Y)^T (Xw - Y) + \lambda_2 w^Y w = Y^T Y - 2Y^T Xw + w^T X^T Xw + \lambda_2 w^T w \tag{2}$$

$$grad_w S = -2X^T Y + 2X^T X\omega + 2\lambda_2 \omega = 0 \tag{3}$$

$$\omega = (X^T X + \lambda_2 E)^{-1} X^T Y \tag{4}$$

Я пробовал преобразовать систему и попытаться решить, её но это сводится к поиску обратной матрицы. И пришел к тому, что нужно попытаться представить как-то матрицу X. Воспользуемся сингулярным разложением: $X = VDU^T$, где D – диагональная матрица из корней собственных значений матриц $X^T X$ и XX^T, U, V – эрмитовы (U составлена из собственных векторов матрицы $X^T X$, V составлена из собственных векторов матрицы XX^T)

Тогда получим:

$$\omega = (UDV^T VD^T U^T + \lambda_2 E)^{-1} UD^T V^T Y \tag{5}$$

$$\omega = (UDD^T U^T + U\lambda_2 EU^T)^{-1} UD^T V^T Y \tag{6}$$

$$\omega = (U^T)^{-1} (D^2 + \lambda_2 E)^{-1} UUDV^T Y \tag{7}$$

$$\omega = U(D^2 + \lambda_2 E)^{-1} DV^T Y \tag{8}$$

$$\omega = \sum_{i=1}^{n} \frac{\sqrt{\lambda_i}}{\lambda_i + \lambda_2} u_i (v_j Y) \tag{9}$$

Т.е. получили, что $w_j(\lambda_2) = \sum\limits_{i=1}^{n} \frac{c_{1,i}}{\lambda_i + \lambda_2}$, где $c_{1,i}$ – какая-то константа полученная от перемножения матриц и соотвествующего коэффициента в u_i и корня из собственного числа.

1.2 II

Теперь рассмотрим для: $S = ||X\omega - Y||_2^2 + \lambda_1||\omega||_1$

$$S = (X\omega - Y)^T(X\omega - Y) + \lambda_1||\omega||_1 \tag{10}$$

$$S = Y^TY - 2Y^TX\omega + \omega^TX^TX\omega + \lambda_1\sum_{i=1}^{n}|w_i| \tag{11}$$

$$grad_wS = -2X^TY + 2X^TX\omega + 2\lambda_1 sign(\omega) = 0 \tag{12}$$

$$\omega = (X^TX)^{-1}(X^TY - \frac{\lambda_1}{2}sign(\omega)) \tag{13}$$

$$\omega = (UD^TV^TVDU^T)^{-1}(UD^TV^T - \frac{\lambda_1}{2}sign(\omega)) \tag{14}$$

$$\omega = (UD^TV^TVDU^T)^{-1}(UD^TV^T - \frac{\lambda_1}{2}sign(\omega)) \tag{15}$$

$$\omega = \sum_{i=1}^{n}(\frac{u_i}{\sqrt{\lambda_i}}(v_i, Y) - \frac{\lambda_1}{2 * \lambda_i}u_i(u_i^T, sign(\omega))) \tag{16}$$

1.3 III

Для случая $S = ||X\omega - Y||_2^2 + \lambda_2||\omega||_2^2 + \lambda_1||\omega||_1$

$$grad_wS = -2X^TY + 2X^TX\omega + 2\lambda_2\omega + \lambda_1 sign(\omega) = 0 \tag{17}$$

$$\omega = (X^TX + \lambda_2E)^{-1}(X^TY - \frac{\lambda_1}{2}sign(\omega)) \tag{18}$$

$$\omega = \sum_{i=1}^{n}\frac{\sqrt{\lambda_i}}{\lambda_i + \lambda_2}\left(u_i(v_j, Y) - \frac{\lambda_1}{2}(u_i, u_i^T)sign(\omega)\right) \tag{19}$$

$$\omega_j = \sum_{i=1}^{n}\frac{\sqrt{\lambda_i}}{\lambda_i + \lambda_2}(u_{i,j}(v_i, Y) - \frac{\lambda_1}{2}u_{i,j}(u_i^T, sign(\omega))) \tag{20}$$