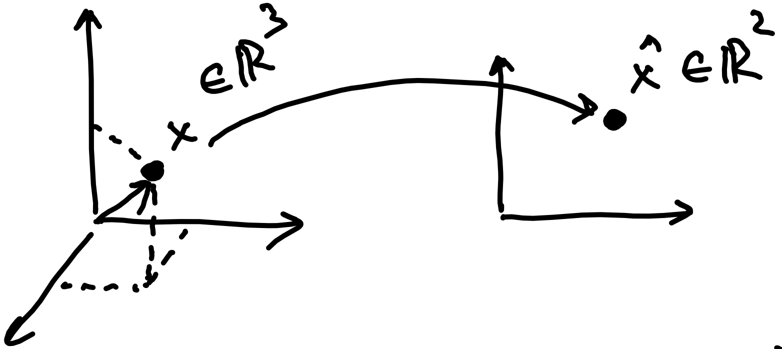
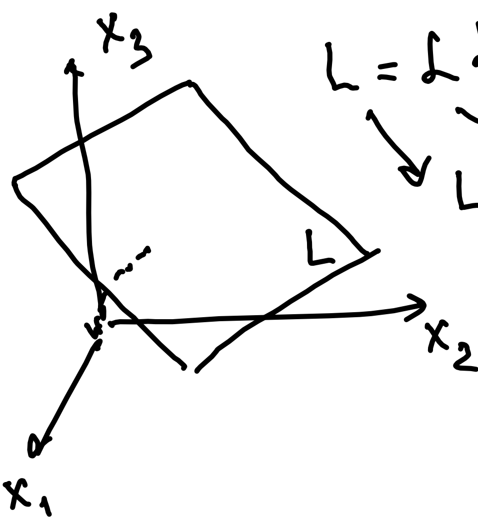
Embeddings.

$$x \in \mathbb{R}^D \longrightarrow \hat{x} \in \mathbb{R}^d$$

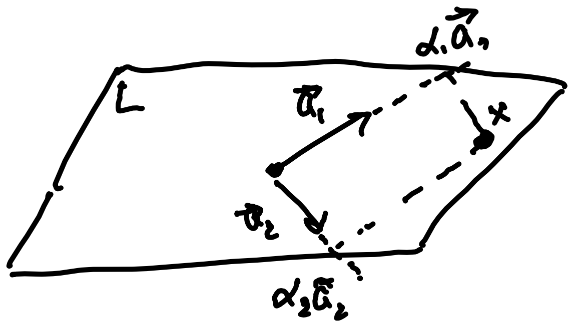


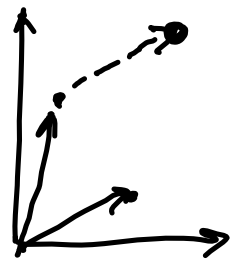


$$L = \mathcal{L}\{\vec{a}_1, \vec{a}_2, \ldots, \vec{a}_d\} \cancel{+\vec{a}}$$

$$L \Longleftarrow A = \underset{D \times d}{\begin{bmatrix} | & & | \\ a_1 & \ldots & a_d \\ | & & | \end{bmatrix}}$$

$$x \in L: \ x = \alpha_1 \vec{a}_1 + \alpha_2 \vec{a}_2 + \ldots + \alpha_d \vec{a}_d$$

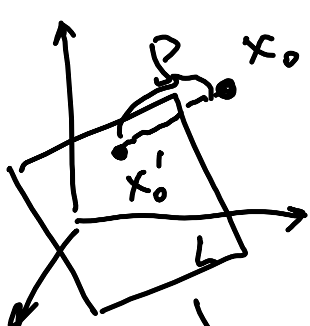




$$\underset{D\times 1}{x} = A\alpha = \underset{D\times d}{\begin{bmatrix} | & & | \\ a_1 & \dots & a_d \\ | & & | \end{bmatrix}} \cdot \underset{d\times 1}{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_d \end{bmatrix}} = \begin{bmatrix} | \\ a_1 \\ | \end{bmatrix} \cdot \alpha_1 + \begin{bmatrix} | \\ a_2 \\ | \end{bmatrix} \cdot \alpha_2 + \dots + \begin{bmatrix} | \\ a_d \\ | \end{bmatrix} \cdot \alpha_d$$

$$\alpha = \underset{d\times 1}{\begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_d \end{bmatrix}}$$

$$x_L = \alpha \in \mathbb{R}^d$$



$L \Rightarrow A$

$$\begin{cases} \frac{1}{2}\|x - x_0\|^2 \to \min\limits_x \\ x = A\alpha \end{cases}$$

$$\Downarrow$$

$$Q = \frac{1}{2}\|A\alpha - x_0\|^2 \to \min_\alpha$$

$$\nabla_x (a^T x) = a, \quad \nabla_x (x^T A x) = (A^T + A) x$$

$$Q = \frac{1}{2} \left(\alpha^T A^T A \alpha - 2 \underbrace{x_0^T A}_{(A^T x_0)^T} \alpha + x_0^T x_0 \right) \to \min_\alpha$$

$$\nabla_\alpha Q = A^T A \alpha - A^T x_0 = 0$$

$$\boxed{\alpha = (\underbrace{A^T A}_{I_d})^{-1} A^T x_0}$$

$$\{ (AB)^T = B^T A^T \}$$

$$A A^T \neq I$$

$$A^T A = \begin{bmatrix} - a_1 - \\ \vdots \\ - a_d - \end{bmatrix} \cdot \begin{bmatrix} | & & | \\ a_1 & \dots & a_d \\ | & & | \end{bmatrix} =$$

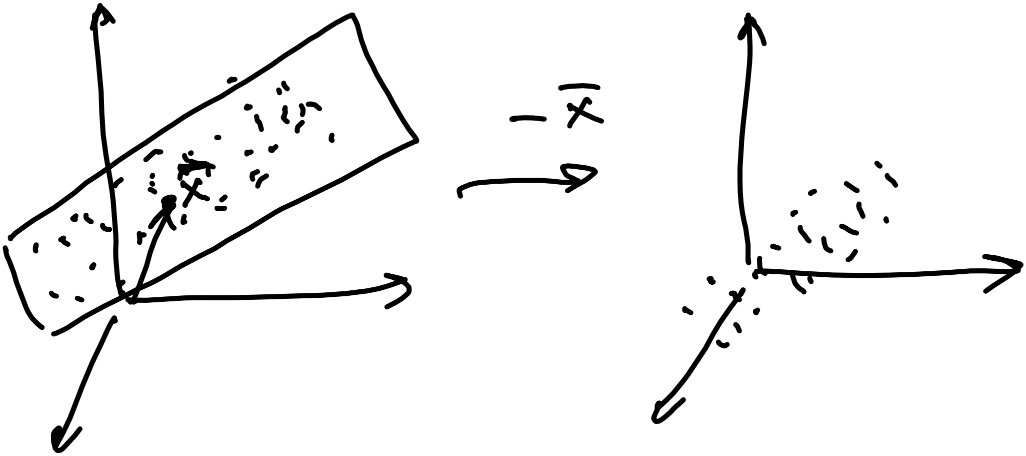
$$= \begin{bmatrix} \langle a_1, a_1 \rangle & \overset{=0}{\langle a_1, a_2 \rangle} & \dots & \overset{=0}{\langle a_1, a_d \rangle} \\ & \langle a_2, a_2 \rangle & & \vdots \\ & & \ddots & \langle a_d, a_d \rangle \end{bmatrix} =$$

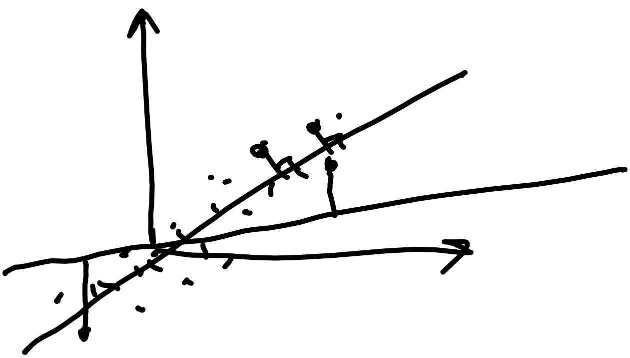
$$= \begin{bmatrix} 1 & & 0 \\ & 1 & \\ 0 & & \ddots & 1 \end{bmatrix} = I_d$$

$$\boxed{\underset{d\times 1}{\alpha} = \underset{d \times D}{A^T} \underset{D\times 1}{x_0}}$$

маломерное представление x_0

$$\operatorname{proj}_L x_0 = \underbrace{\underset{D\times d}{A} \underset{d\times D}{A^T} \underset{D\times 1}{x_0}}_{D\times 1}$$





$$Q = \frac{1}{n}\sum_{i=1}^{n} \rho(x_i, \underset{\substack{\downarrow\\ A}}{L}) \to \min_{L}$$

$$\frac{1}{n}\sum_{i=1}^{n} \| x_i - AA^T x_i \|^2 \to \min_{A = a_1, a_2, \dots, a_d}$$

$$\frac{1}{n}\sum_{i=1}^{n} (x_i^T - x_i^T AA^T)(x_i - AA^T x_i) \to \min$$

$$\frac{1}{n}\sum_{i=1}^{n} x_i^T \underbrace{(I - AA^T)(I - AA^T)}_{I - AA^T} x_i \ominus$$

$$\left\{ (I - AA^T)(I - AA^T) = I - AA^T - AA^T + A \underbrace{A^T A}_{I_d} A^T = I - 2AA^T + AA^T = I - AA^T \right\}$$

$$\ominus \frac{1}{n}\sum_{i=1}^{n} x_i^T (I - AA^T) x_i \to \min_A$$

$$\cancel{\frac{1}{n}\sum_i x_i^T x_i} - \frac{1}{n}\sum_i \underbrace{\underbrace{x_i^T}_{(A^T x_i)^T} A A^T x_i}_{\|A^T x_i\|^2} \to \min_A$$

$$Q = \frac{1}{n}\sum_{i=1}^{n} \|A^T x_i\|^2 \to \max_A$$

$$A^T x_i = \underset{d \times D}{\begin{bmatrix} - & a_1 & - \\ & \vdots & \\ - & a_d & - \end{bmatrix}} \cdot \underset{D \times 1}{\begin{bmatrix} | \\ x_i \\ | \end{bmatrix}} = \underset{d \times 1}{\begin{bmatrix} \langle a_1, x_i \rangle \\ \langle a_2, x_i \rangle \\ \vdots \\ \langle a_d, x_i \rangle \end{bmatrix}}$$

$$\|A^T x_i\|^2 = \langle a_1, x_i\rangle^2 + \langle a_2, x_i\rangle^2 + \ldots + \langle a_d, x_i\rangle^2 =$$

$$= \sum_{k=1}^{d} \langle a_k, x_i\rangle^2$$

$$Q = \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{d} \langle a_k, x_i\rangle^2 \to \max_{a_1,\ldots,a_d}$$

$$\langle a_k, x_i\rangle = a_k^T x_i = x_i^T a_k$$

$$\langle a_k, x_i\rangle^2 = \langle a_k, x_i\rangle \cdot \langle a_k, x_i\rangle =$$

$$= a_k^T x_i \cdot x_i^T a_k$$

$$Q = \frac{1}{n}\sum_{i=1}^{n}\sum_{k=1}^{d} a_k^T \underbrace{x_i \cdot x_i^T} a_k =$$

$$= \sum_{k=1}^{d} a_k^T \underbrace{\left(\frac{1}{n}\sum_{i=1}^{n} x_i x_i^T\right)}_{S} \cdot a_k \to \max_{a_1,\ldots,a_d}$$

$$\begin{cases} \sum_{k=1}^{d} a_k^T S a_k \to \max\limits_{a_1,\ldots,a_d} \Rightarrow \text{распадается на } d \text{ подзадач.} \\ \|a_k\|^2 = 1 \quad k = 1,\ldots,d \\ \langle a_i, a_j\rangle = 0 \quad \forall i \neq j \to \text{не нужно.} \end{cases}$$

$$\sim \begin{cases} a_k^T S a_k \to \max\limits_{a_k} \\ \|a_k\|^2 = 1 \end{cases}$$

$$L = a_k^T S a_k + \mu (a_k^T a_k - 1)$$

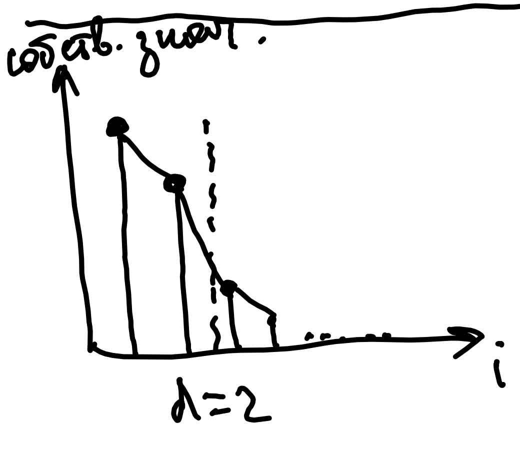
$$\nabla_{a_k} L = 2 S a_k + 2 \mu a_k = 0.$$

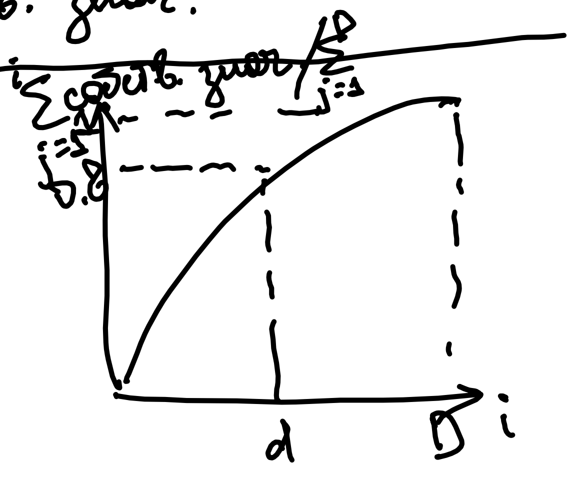
$$S a_k = -\mu a_k \Rightarrow \boxed{S a_k = \gamma a_k}$$

a_k – собственный вектор для S,

$$a_k^T S a_k = a_k^T \cdot \gamma a_k = \gamma \cdot \underbrace{a_k^T a_k}_{1} = \gamma \to \max$$

соотв. макс. собств. знач.





$\boxed{\text{PCA}}$, $a_k \in \mathbb{R}^D$

$$\alpha_x \in \mathbb{R}^d = \underset{d \times D}{A^T} x$$

Семинар.

1. $$\begin{cases} a_k^T \Sigma a_k \to \max\limits_{a_k} \\ \|a_k\|^2 = 1 \end{cases}$$

2. $$Q = \frac{1}{n} \sum_{i=1}^{n} \log\left(1 + \exp\left(\underbrace{-y_i \langle w, x_i \rangle}_{\text{Margin}}\right)\right) + \lambda \sum_{k=1}^{d} |w_k|$$

3. $$G(R) = \sum_{i=1}^{K} p_i (1 - p_i)$$

$$p_k = \frac{1}{|R|} \cdot \sum_{i:(x_i, y_i) \in R} [y_i = k]$$

$$Q = G(R) - \frac{|R_r|}{|R|} \cdot G(R_r) - \frac{|R_l|}{|R|} \cdot G(R_l)$$