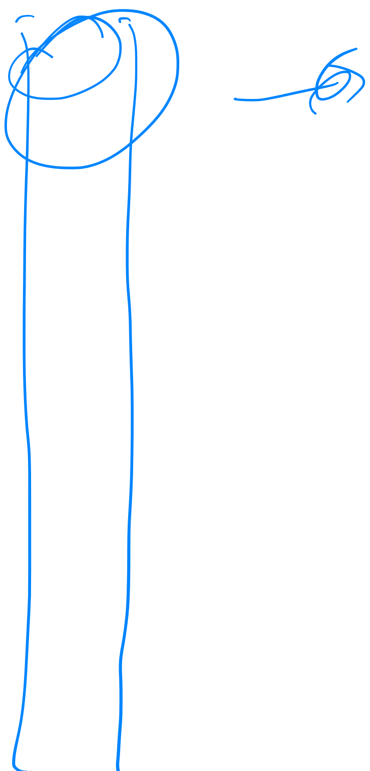
Random Projection

$d = 1M$

$d \to k$



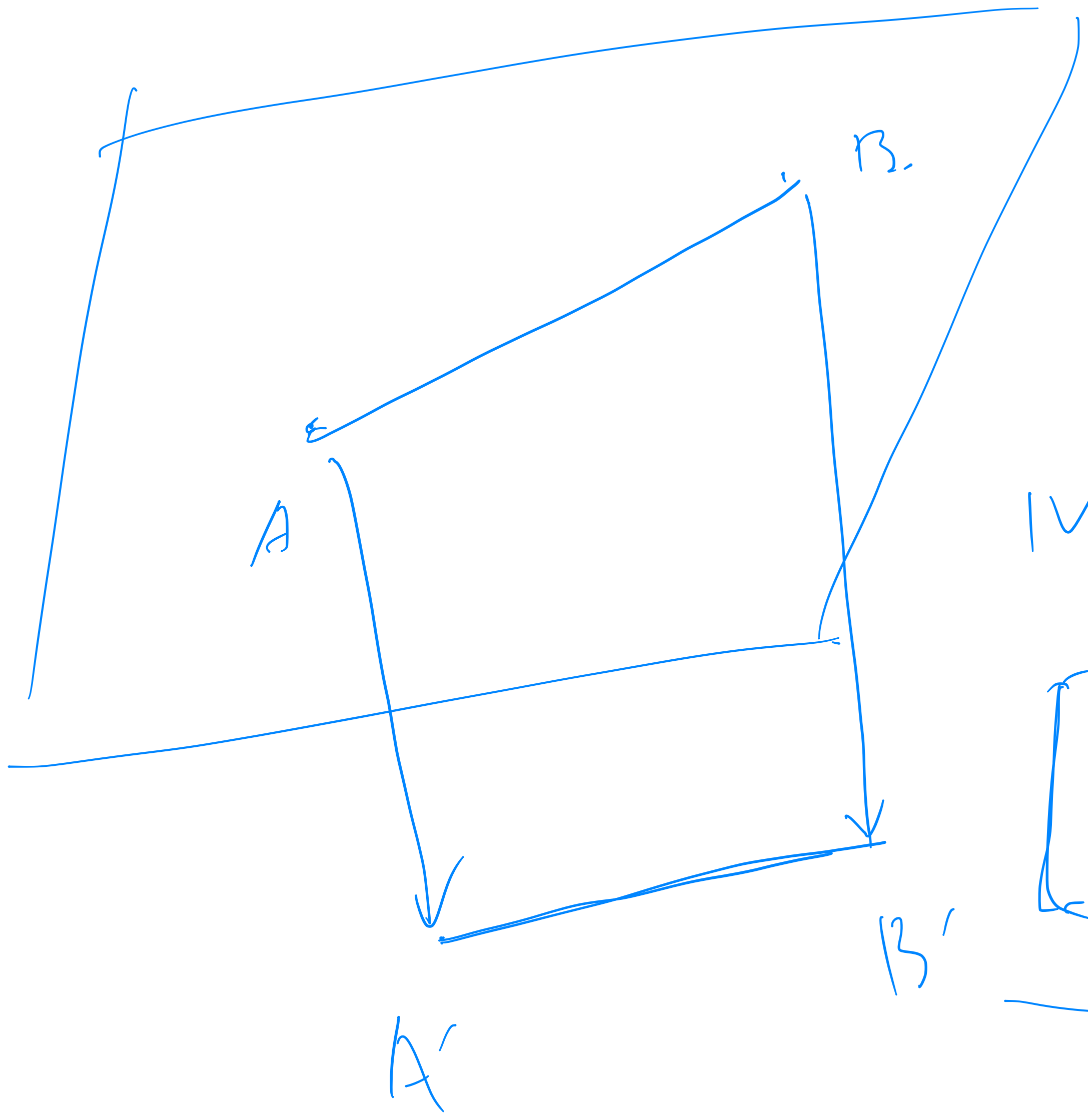
v $\mathbb{R}^d$ $\qquad\qquad$ $\mathbb{R}^k$.

M : $\mathbb{R}^{k \times d}$ (freshman).

M v $\longrightarrow$ $\underline{M \cdot v}$

k - dim.

$\mathbb{R}^d$.



$M_1 = \cancel{D_{k \times d}}$

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

$$\| A' - B' \| \approx \| A - B \|$$

$\epsilon \in (0, 1)$

$$\boxed{(1-\epsilon) \cdot \|A - B\| \le \|A' - B'\| \le (1+\epsilon) \cdot \|A - B\|}$$

Goal.

$\epsilon = 0.01$

$M_{ij} \sim N(0, 1)$

$\boxed{M : k \times d}$ $\{v_1 \; --- \; v_n\}$

Step 1. $M_{ij} \sim N(0, 1)$

Step 2. $v_i \to M \cdot v_i$

$\mathbb{R}^d \qquad \mathbb{R}^k$

$d = 10^{1000}$

$k \geq \frac{\log n}{\epsilon^2}$

$\boxed{k = 10000 \log n}$

when $\epsilon = 0.01$

7:35

$\forall\, i, j \in \{1, \dots, n\}$

$$(1-\epsilon)\cdot\|v_i - v_j\| \le \|M v_i - M v_j\| \le (1+\epsilon)\cdot\|v_i - v_j\|$$

成立

$$v \qquad = M\cdot(v_i - v_j)$$

$v \in \mathbb{R}^d$.

$$(1-\epsilon)\cdot\|v\|^2 \le \|Mv\|^2 \le (1+\epsilon)\cdot\|v\|^2$$

$$v^T v$$

$$\|Mv\|_2^2 = v^T M^T M v$$

$$= v^T (M^T M) v$$

① $\boxed{E[v^T (M^T M) v] = v^T v}$

② Exp $\rightarrow$ high prob.

$$E\left[\underline{v^T M^T M v}\right]$$

$$= v^T \cdot \underline{E\left[M^T M\right]} \cdot v \quad (\text{not hard})$$

$$\approx I$$

$$M = \begin{bmatrix} \uparrow & & \uparrow \\ m_1 & \cdots & m_d \\ \downarrow & & \downarrow \end{bmatrix}_{k \times d}$$

$$M^T = \begin{bmatrix} m_1^T \\ \vdots \\ m_d^T \end{bmatrix}$$

$$M^T M = \begin{bmatrix} \leftarrow m_1^T \rightarrow \\ \vdots \\ m_d^T \end{bmatrix} \begin{bmatrix} m_1 & \cdots & m_d \end{bmatrix}$$

$$\underbrace{M^T M}_{d \times d} = \begin{bmatrix} \leftarrow m_1^T \rightarrow \\ \vdots \\ m_d^T \end{bmatrix} \begin{bmatrix} m_1 & \cdots & m_d \end{bmatrix}$$

$$= \begin{bmatrix} m_1^T m_1 & m_1^T m_2 & \cdots & m_1^T m_d \\ m_2^T m_1 & & & \\ & & & \\ & & & m_d^T m_d \end{bmatrix}$$

$$E[m_i^T m_j] \qquad 1 \le i, j \le d.$$

$$E[a^T b] \qquad i \ne j \quad E[m_i^T m_j] = 0$$

$$= E\left[\sum_{i=1}^{k} a_i b_i\right]$$

$$= \sum_{i=1}^{k} E[a_i b_i] = \underbrace{\sum E[a_i] \cdot E[b_i]}_{= 0}$$

~~$i \neq j$~~

$\Rightarrow \quad a = b$

$a_i \sim N(0,1)$

$Var[a_i] = 1$

$$E[a^T \cdot a]$$

$$= E\left[\sum_{i=1}^{k} a_i^2\right] \qquad = E[a_i^2] - 0$$

$$= \sum_{i=1}^{k} \underbrace{E[a_i^2]}_{=1}$$

$$= k$$

$\Rightarrow \quad i = j$

$$E[m_i^T m_j] = k$$

$$\Rightarrow E[M^T M] = k \cdot I$$

$$\tilde{M} = M / \sqrt{k} \qquad E[\tilde{M}^T \cdot \tilde{M}] = I$$

$v_1 \, \ldots \, v_n \in \mathbb{R}^d.$

$$d \rightarrow \boxed{k} \geq \frac{\log n}{\epsilon^2}$$

① $\tilde{M}$. $\tilde{M}_{ij} \sim N(0, 1/k)$

② $v_i \rightarrow \tilde{M} \cdot v_i$

$$E\left[\| \tilde{M} \cdot v \|^2 \right] = \| v \|^2$$

Let
$v = v_i - v_j$

$$E\left[\| \underbrace{\tilde{M} (v_i - v_j)} \|^2 \right]$$

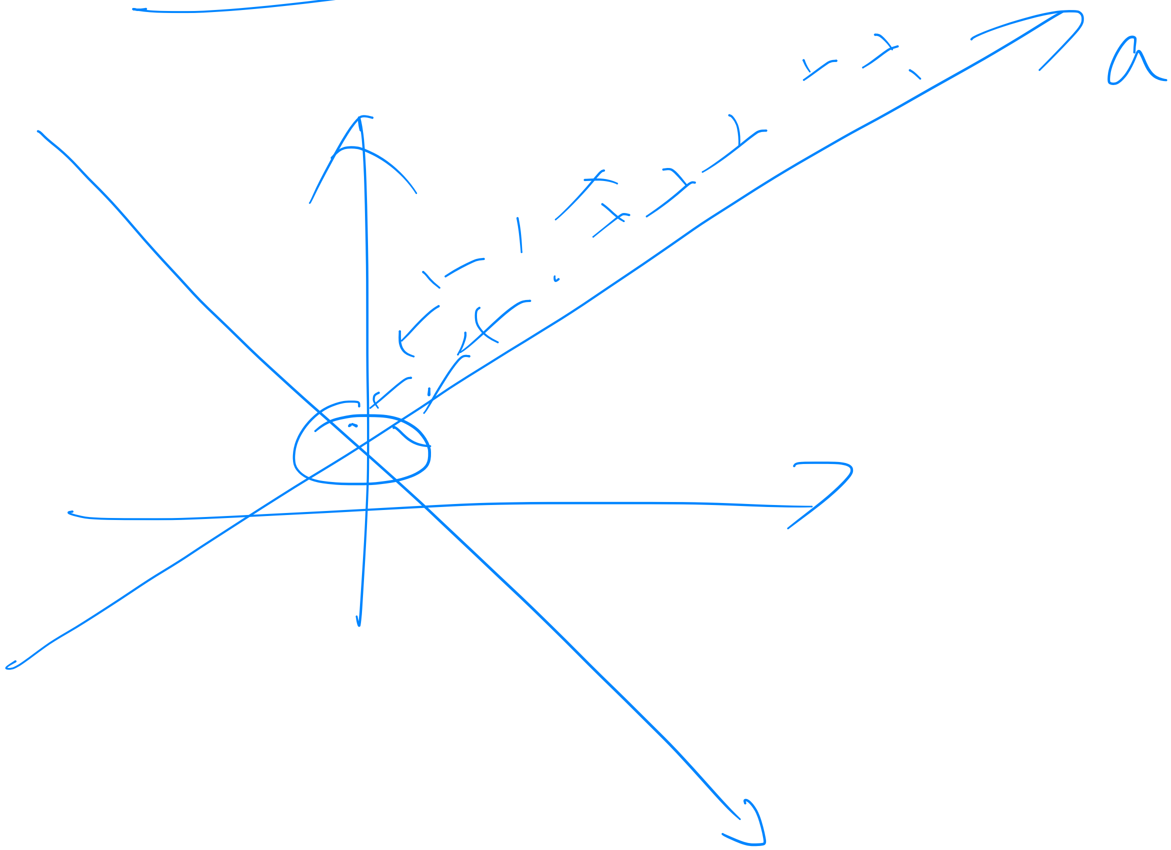
$$= \| v_i - v_j \|^2$$

$\tilde{M}$ as sparse

$\{+1, -1, 0\}$

$$\tilde{M}_{ij} = \begin{cases} +1 & \text{w.p. } 1/6 \\ -1 & 1/6 \\ 0 & 2/3. \end{cases}$$

Principal Component Analysis



$$X = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 6 & 8 & 10 \end{bmatrix}$$

$$X = \begin{bmatrix} 1.001 & 2.001 & 3.01 & 3.999 & 4.99 \\ 2 & 4 & 6 & 8 & 10 \end{bmatrix}$$

a. find low-rank M.

b. $M \simeq X$

$$\boxed{\begin{cases} \min_{M} \|M - X\|_F \\ \text{s.t. } \operatorname{rank}(M) \le r \end{cases}}$$

non-convex

$\Rightarrow$ 1990's.

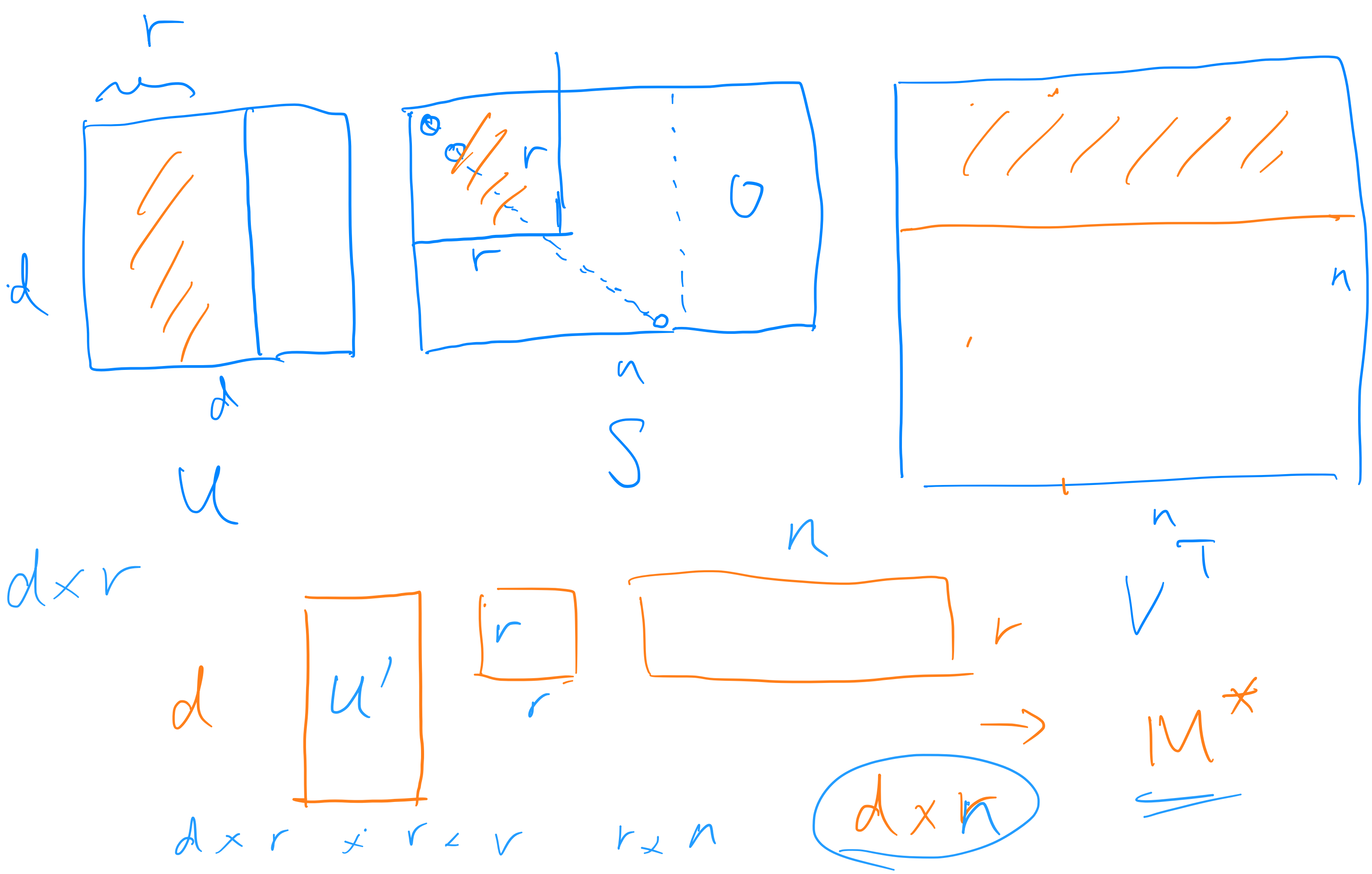
singular value decomposition

$$X \to \boxed{SVD} \to (U, S, V).$$

① $X = USV^T$ $\quad X: d \times n$

② $U^TU = I, \quad V^TV = I$

③ S is diagonal.



dim - reduction.

$$(M^*)^T X \to \underline{r \times n}.$$

$$(U')^T \cdot \underline{M^*}$$

$d \times n$