Recap

$$\varphi : V \longrightarrow W$$

$v_1, \ldots, v_r, v_{r+1} \ldots v_n$ $\qquad$ $w_1, \ldots, w_r, w_{r+1} \ldots w_m$ $\quad \left(w_i = \dfrac{\varphi(v_i)}{\sigma_i}\right)$

$$\varphi = \sum_{i=1}^{r} \underbrace{\sigma_i}_{\text{singular values}} \cdot \underbrace{|w_i\rangle}_{\text{left singular vectors}} \underbrace{\langle v_i|}_{\text{right singular vectors}}$$

$A \in \mathbb{C}^{m \times n}$ $\qquad$ $A = \sum_{i=1}^{r} \sigma_i \cdot w_i v_i^*$

$$A = \begin{bmatrix} | & & | & & | \\ w_1 & \cdots & w_r & \cdots & w_m \\ | & & | & & | \end{bmatrix} \left[\begin{array}{ccc|c} \sigma_1 & & 0 & \\ & \ddots & & 0 \\ 0 & & \sigma_r & \\ \hline & 0 & & 0 \end{array}\right] \begin{bmatrix} — & v_1^* & — \\ & \vdots & \\ — & v_r^* & — \\ & & \\ — & v_n^* & — \end{bmatrix}$$

Application : Low-rank approximation

Given : $A \in \mathbb{R}^{m \times n}$, $\text{rank}(A) = r$ (say)

Find : $A_k \in \mathbb{R}^{m \times n}$, $\text{rank}(A_k) \leq k$

s.t. is "close to" A $\Leftrightarrow$ norm of $A - A_k$ is "small"

Matrix norms

Spectral Norm $\|M\|_2 = \max_{v \neq 0} \dfrac{\|Mv\|_2}{\|v\|_2} = \max_{v \neq 0} \dfrac{\sqrt{\langle Mv, Mv \rangle}}{\sqrt{\langle v, v \rangle}}$

$$= \sqrt{\max_{v \neq 0} \frac{\langle v, M^* M v \rangle}{\langle v, v \rangle}}$$

Frobenius Norm $\|M\|_F = \left(\sum_{i,j} |M_{ij}|^2 \right)^{1/2}$

Low-rank approximation via SVD

Given $A = \sum_{i=1}^{r} \sigma_i \cdot u_i v_i^*$ $\qquad A_k = \sum_{i=1}^{k} \sigma_i \, u_i \cdot v_i^*$

$(\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r > 0)$

- $\| A - A_k \|_2 = \sigma_{k+1}$

- $\forall B$, $\text{rank}(B) \leq k$ $\quad \| A - B \|_2 \geq \sigma_{k+1}$

▸ $\|A - A_k\|_2 = \sigma_{k+1}$

<u>Proof:</u> For $k = 0$, $\|A\|_2 = \sigma_1$ since $\sigma_1^2 = \max\limits_{v \neq 0} \dfrac{\langle v, A^* A v\rangle}{\langle v, v\rangle}$

$A - A_k = \sum\limits_{i=k+1}^{r} \sigma_i w_i v_i^*$ has the same singular vectors

$$(A - A_k)^*(A - A_k)\, v_j = (A - A_k)^* \left(\sum_{i=k+1}^{r} \sigma_i w_i \langle v_i, v_j\rangle \right)$$

$$= \begin{cases} (A - A_k)^* \sigma_j w_j & \text{if } j \in [k+1, r] \\ 0 & \text{otherwise} \end{cases}$$

$$= \begin{cases} \sigma_j^2 \cdot v_j & \text{if } j \in [k+1, r] \\ 0 & \text{otherwise.} \end{cases}$$

$\therefore$ $v_1 \ldots v_n$ is still an orthonormal eigenbasis for $(A - A_k)^*(A - A_k)$

$\therefore$ $\|A - A_k\|_2 = \sigma_{k+1}$ (largest singular value)

- $\forall B$, $\operatorname{rank}(B) \leq k$ $\quad \|A-B\|_2 \geq \sigma_{k+1}$

Proof: By Ex below, $\dim(\operatorname{Span}\{v_1 \dots v_{k+1}\} \cap \ker(B)) \geq 1$

$$\exists\ z = \sum_{i=1}^{k+1} c_i \cdot v_i \in \operatorname{Span}\{v_1 \dots v_{k+1}\} \cap \ker(B)$$

$$\frac{\|(A-B)z\|_2^2}{\|z\|_2^2} = \frac{\left\| \sum_{i=1}^{k+1} c_i \cdot \sigma_i \cdot w_i \right\|_2^2}{\left\| \sum c_i \cdot v_i \right\|_2^2}$$

$$= \frac{\sum_{i=1}^{k+1} |c_i|^2 \cdot \sigma_i^2}{\sum_{i=1}^{k+1} |c_i|^2} \geq \frac{\sum_{i=1}^{k+1} |c_i|^2 \cdot \sigma_k^2}{\sum_{i=1}^{k} |c_i|^2}$$

Ex: S_1, S_2 subspaces of finite dimensional V. Then,

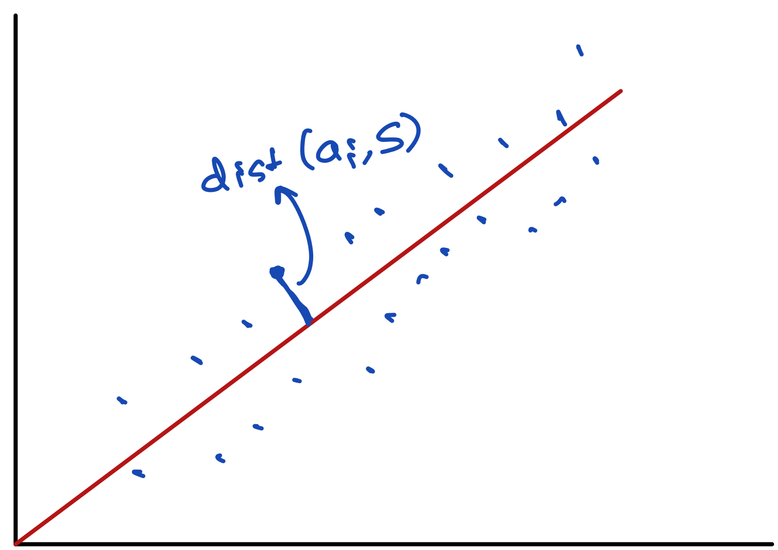
$$\dim(S_1 \cap S_2) \geq \dim(S_1) + \dim(S_2) - \dim(V)$$

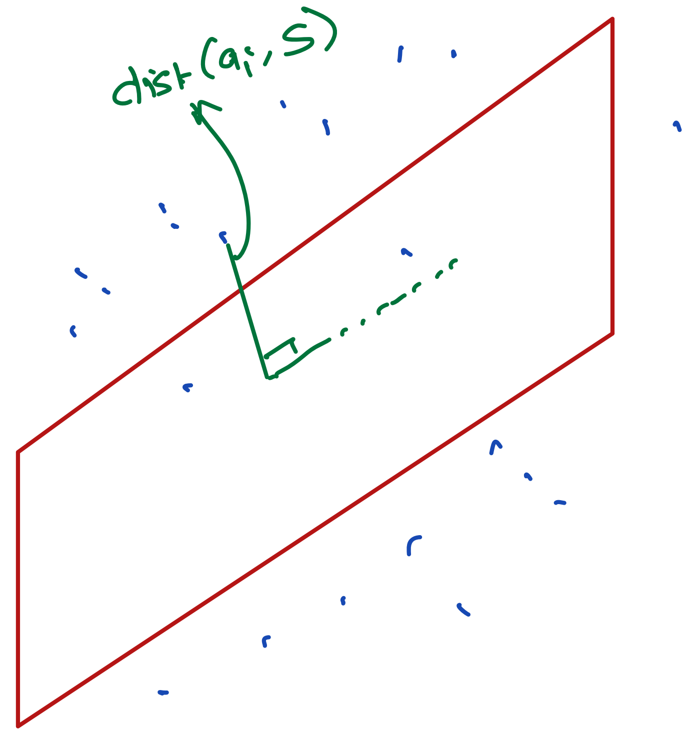
Least Squares Approximation (dimension reduction / PCA)

Given : $a_1, \ldots, a_n \in \mathbb{R}^d$ (with standard inner product)

Find : Subspace $S \subseteq \mathbb{R}^d$, $\dim(S) \le k$

minimizing $\sum_{i=1}^{n} (\mathrm{dist}(a_i, S))^2$





Calculating distance to a subspace S

Given orthonormal basis $u_1, \ldots, u_k$ for $S \subseteq V$ $(V = \mathbb{R}^d)$

$$\operatorname{dist}(v, S)^2 = \min_{u \in S} \|u - v\|^2 = \|v\|^2 - \sum_{j=1}^{k} \langle v, u_j \rangle^2$$

$$\|v\|^2 = \sum_{j=1}^{d} c_j^2 \qquad \sum_{j=1}^{k} \langle v, u_j \rangle^2 = \sum_{j=1}^{k} c_j^2$$

<u>Proof:</u> Complete $u_1, \ldots, u_k, u_{k+1}, \ldots, u_d$ to o.n. basis for V

For $v = \sum_{j=1}^{d} c_j u_j$ $\qquad u \in S$, $u = \sum_{j=1}^{k} b_j u_j$

$$\|v - u\|^2 = \Big\| \sum_{j=1}^{k} (c_j - b_j) \cdot u_j + \sum_{j>k} c_j \cdot u_j \Big\|^2$$

$$= \sum_{j=1}^{k} (c_j - b_j)^2 + \boxed{\sum_{j>k} c_j^2}$$

minimized for $u = \sum_{j=1}^{k} c_j u_k$

Restating the problem

Given : $a_1, \ldots, a_n \in \mathbb{R}^d$ (with standard inner product)

Find : Orthonormal $u_1, \ldots, u_k$ $(S = \mathrm{Span}\{u_1, \ldots, u_k\})$

minimizing $\sum_{i=1}^{n} \left(\|a_i\|^2 - \sum_{j=1}^{k} \langle a_i, u_j \rangle^2 \right)$

maximizing $\sum_{j=1}^{k} \sum_{i=1}^{n} \langle a_i, u_j \rangle^2 = \sum_{j=1}^{k} \|A u_j\|^2$

$$A = \begin{bmatrix} \text{---} & a_1^T & \text{---} \\ & \vdots & \\ \text{---} & a_n^T & \text{---} \end{bmatrix}$$

For $\begin{pmatrix} | \\ u \\ | \end{pmatrix}$ $Au = \begin{pmatrix} \\ \langle a_i, u \rangle \\ \\ \end{pmatrix}$ $\|Au\|^2 = \sum \langle a_i, u \rangle^2$

Formulating as a matrix problem

Given : $A \in \mathbb{R}^{n \times d}$

Find : Orthonormal $u_1 \ldots u_k \in \mathbb{R}^d$

maximizing $\|Au_1\|^2 + \cdots + \|Au_k\|^2$

Solution: SVD $\quad A = \sum_{i=1}^{r} \sigma_i w_i v_i^*$

$\sigma_1^2 \geq \cdots \geq \sigma_r^2 > 0$

$v_1 \ldots v_r \in \mathbb{R}^d$

$w_1 \ldots w_r \in \mathbb{R}^n$

For $k \leq r$, take

$u_1 = v_1, \ldots, u_k = v_k$

Ex: Problem is trivial if $k > r$ (all points lie on subspace of dimension $r < k$)

Optimality

- Let $v_1, \dots, v_r \in \mathbb{R}^d$ be the right singular vectors of A. corresponding to singular values $\sigma_1 \geq \dots \geq \sigma_r > 0$. Then $\forall k \leq r$, $\forall$ orthonormal $u_1 \dots u_k$

$$\|Au_1\|^2 + \dots + \|Au_k\|^2 \leq \|Av_1\|^2 + \dots + \|Av_k\|^2$$

Proof: Induction on k.

$k=1$ $\quad \forall u_1 \quad \|Au_1\|^2 \leq \|Av_1\|^2$

$$\max_{\|u_1\|=1} \langle Au_1, Au_1 \rangle = \sigma_1^2 = \|Av_1\|^2$$

Induction: Assume for all orthonormal $u_1 \dots u_{k-1}$

$$\|Au_1\|^2 + \dots + \|Au_{k-1}\|^2 \leq \|Av_1\|^2 + \dots + \|Av_{k-1}\|^2$$

Say $u_k \in V_{k-1}^{\perp} = \{u \in V \mid \langle u, v_1 \rangle = \dots = \langle u, v_{k-1} \rangle = 0\}$

Then

$$\|Au_k\|^2 \leq \max_{\substack{u \in V_{k-1}^{\perp} \\ \|u\|=1}} \|Au\|^2 = \max_{\substack{u \in V_{k-1}^{\perp} \\ \|u\|=1}} \langle u, A^* A u \rangle$$

$$= \;?$$

Let $u = \sum_{i=k}^{d} c_i v_i$ with $\sum c_i^2 = 1$. $\langle u, A^* A u \rangle = \sum_{i=k}^{d} c_i^2 \cdot \sigma_i^2 \leq \sigma_k^2$

$$= \|Av_k\|^2$$

Justifying the assumption $u_k \in V_{k-1}^\perp$

- Given orthonormal $u_1, \ldots, u_k$, $\exists$ orthonormal $u_1', \ldots, u_k'$ s.t.
  - $u_k' \in V_{k-1}^\perp$
  - $\text{Span}\{u_1, \ldots, u_k\} = \text{Span}\{u_1', \ldots, u_k'\}$ S

    $\Downarrow$
  - $\|Au_1\|^2 + \cdots + \|Au_k\|^2 = \|Au_1'\|^2 + \cdots + \|Au_k'\|^2$

Proof:

$$\dim(S \cap V_{k-1}^\perp) \geq k + (n-k+1) - n \geq 1$$