Subspaces in Machine Learning

Objectives

- Define subspace
- Establish centrality of subspaces in machine learning
- Introduce dimension of low-rank approximations

"Subspaces" play a key role in machine learning

- Classification/data modeling : $\underline{A}\,\underline{w} = \underline{d}$

$$\underline{A}\,\underline{w} = \left[\underline{a}_1\ \underline{a}_2 \cdots \underline{a}_m\right]\begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_m \end{bmatrix} = \sum_{i=1}^{m} \underline{a}_i w_i$$

Span$\{\underline{a}_1, \underline{a}_2, \cdots \underline{a}_m\}$

- Modeling matrix data $\quad \hat{\underline{R}} = \underline{T}\,\underline{S}$

$$\hat{\underline{R}} = \left[\underline{t}_1\ \underline{t}_2 \cdots \underline{t}_m\right]\begin{bmatrix} \underline{s}_1^T \\ \underline{s}_2^T \\ \vdots \\ \underline{s}_m^T \end{bmatrix} = \sum_{i=1}^{m} \underline{t}_i \underline{s}_i^T \quad \text{or} \quad \hat{\underline{R}} = \left[\hat{\underline{r}}_1\ \hat{\underline{r}}_2 \cdots \hat{\underline{r}}_k\right] \Rightarrow \hat{\underline{r}}_j = \sum_{i=1}^{m} \underline{t}_i s_{ij}$$

Span$\{\underline{t}_1, \underline{t}_2, \cdots, \underline{t}_m\}$

Span$\{\underline{v}_1, \underline{v}_2, \cdots \underline{v}_m\}$ is a subspace

Formally, a **subspace** $S \subseteq \mathbb{R}^N$ (set of N-dim points) satisfies

i) $\underline{0} \in S$ (contains origin)

ii) if $\underline{f}, \underline{g} \in S$, then $\underline{f} + \underline{g} \in S$ (closed under addition)

iii) if $\underline{f} \in S$, then $\alpha \underline{f} \in S$ (closed under scalar mults)

Examples: $S = \{(x,y,z) \mid z = 0\}$ i) $(0,0,0) \in S$ ✓

ii) $(f_x, f_y, 0) + (g_x, g_y, 0) = (f_x + g_x, f_y + g_y, 0) \in S$ ✓

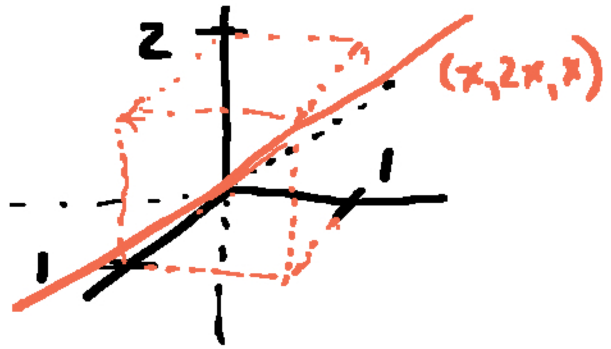
iii) $\alpha(f_x, f_y, 0) = (\alpha f_x, \alpha f_y, 0) \in S$ ✓ Subspace (x-y plane)

- $S = \{(x,y,z) \mid z = 1\}$ i) $(0,0,0) \notin S$ not a subspace

- $S = \{(x, 2x, x)\}$ (line in 3d) i) $(0,0,0) \in S$ ✓

ii) $(f, 2f, f) + (g, 2g, g) = (f+g, 2(f+g), f+g) \in S$ ✓

iii) $\alpha(f, 2f, f) = (\alpha f, 2\alpha f, \alpha f) \in S$ ✓ Subspace (line)



Consider $S \subseteq \mathbb{R}^N$, $\left\{ [\underline{v}_1\ \underline{v}_2 \cdots \underline{v}_M] \begin{bmatrix} w_1 \\ w_2 \\ \vdots \\ w_M \end{bmatrix} = \underline{V}\,\underline{w} \text{ for } \underline{w} \in \mathbb{R}^M \right\}$

$\underline{V} : N \times M$

i) if $\underline{w} = \underline{0}$, $\underline{V}\,\underline{w} = \underline{0} \in S$

ii) Let $\underline{f} = \underline{V}\,\underline{w}_f$, $\underline{g} = \underline{V}\,\underline{w}_g$, then $\underline{f} + \underline{g} = \underline{V}\,\underline{w}_f + \underline{V}\,\underline{w}_g = \underline{V}(\underline{w}_f + \underline{w}_g) \in S$

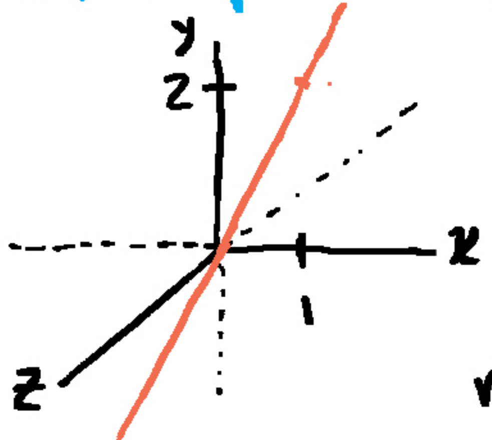
iii) Let $\underline{f} = \underline{V}\,\underline{w}_f$, then $\alpha \underline{f} = \alpha \underline{V}\,\underline{w}_f = \underline{V}(\alpha \underline{w}_f) \in S$

Subspace!

Dimension of $\{\underline{V}\,\underline{w}\}$ is $\text{rank}(\underline{V})$

Examples:

$\underline{V} = \begin{bmatrix} 1 \\ 2 \\ 0 \end{bmatrix}$



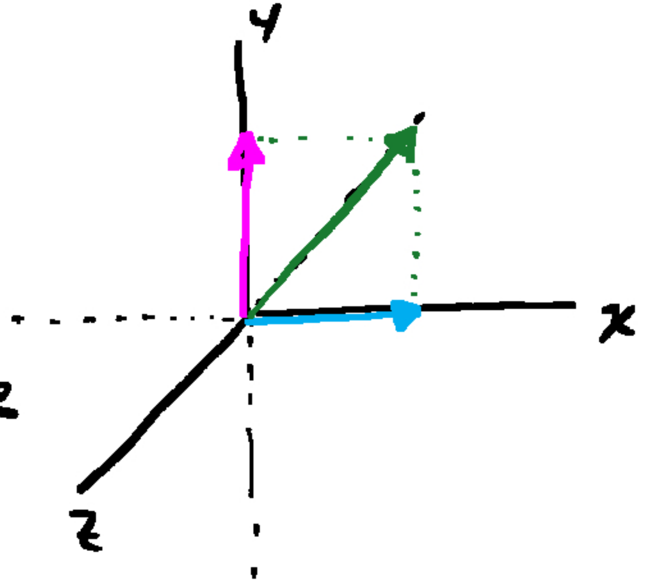
$\underline{V}\,\underline{w}$: line in x-y plane

$\text{rank}(\underline{V}) = 1$

$\underline{V} = \begin{bmatrix} 1 & 0 & 1 \\ 1 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}$

$\underline{V}\,\underline{w}$: x-y plane

$\text{rank}(V) = 2$



In general $S=\{\underline{V}\,\underline{w}\}$ is a $K=\text{rank}\,\underline{V}$ dim hyperplane in $\mathbb{R}^N$
(origin)

What about $\hat{\underline{R}} = \underline{T}\,\underline{S}$?

- In general $\text{rank}(\underline{T}\,\underline{S}) \le \min\{\text{rank}(\underline{T}), \text{rank}(\underline{S})\}$
(proof in notes)
- Special case: $\underline{T}: N\times M$, $\text{rank}(\underline{T})=M$
$\underline{S}: M\times K$, $\text{rank}(\underline{S})=M$ iff

$\text{rank}(\hat{\underline{R}}) = M$ (proof in notes)

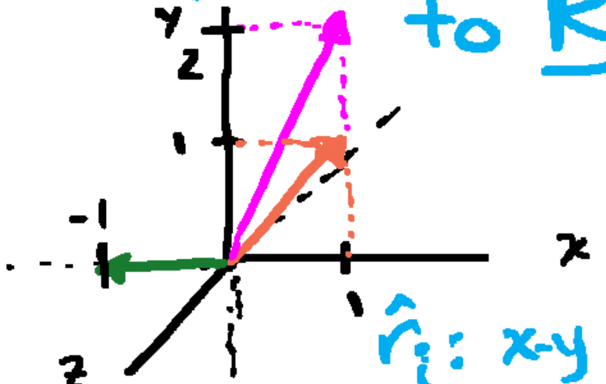
$$\underline{R} \approx \hat{\underline{R}}, \quad \hat{\underline{R}} = \underline{T}\,\underline{S} = \sum_{i=1}^{M} \underline{t}_i \underline{s}_i^T$$

rank-M approximation to $\underline{R}$

$\hat{\underline{R}} = [\hat{\underline{r}}_1, \hat{\underline{r}}_2, \cdots, \hat{\underline{r}}_K]$

$$\hat{\underline{R}} = \underbrace{\begin{bmatrix} 1 & 0 \\ 1 & 1 \\ 0 & 0 \end{bmatrix}}_{\text{rank } 2} \overbrace{\begin{bmatrix} 1 & 1 & -1 \\ 0 & 1 & 1 \end{bmatrix}}^{\text{rank } 2} = \begin{bmatrix} 1 & 1 & -1 \\ 1 & 2 & 0 \\ 0 & 0 & 0 \end{bmatrix}$$

$\hat{\underline{r}}_i$ lie in M-dimensional subspace



$\hat{\underline{r}}_i$: x-y plane