Representing Data with Bases

Objectives

- Introduce bases as building blocks for subspaces
- Introduce example uses of bases
- Define orthonormal bases

A subspace can be described as the span of a set of vectors.

$$S = \left\{ \underline{x} : \underline{x} = \sum_{i=1}^{M} \underline{v}_i w_i \,,\ w_i \in \mathbb{R},\ i=1,2,\cdots M \right\}$$
$$= \text{span}\{\underline{v}_i\}$$

1) $\underline{v}_i$ arbitrary - hard computing if linear dep.

2) $\underline{v}_i$ linearly independent

3) $\underline{v}_i$ orthonormal - easiest computing

Basis

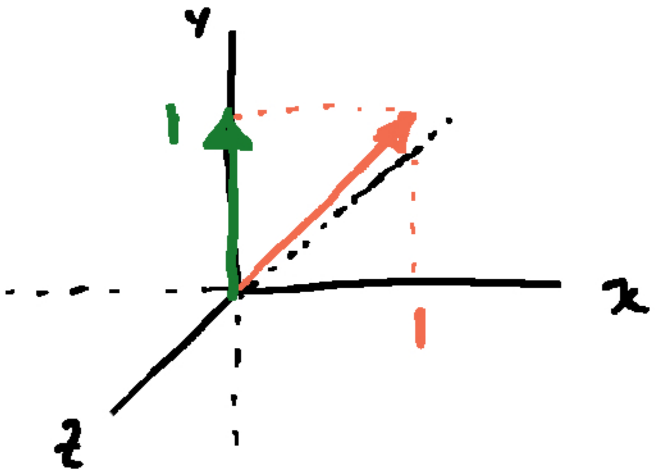
- unique relationship between $\underline{x}$ and w_i in 2), 3)

Orthonormal: $\underline{v}_i$ are orthognal to each other and unit length

$$\underline{v}_i^T \underline{v}_i = 1 \,,\quad \underline{v}_i^T \underline{v}_j = 0 \quad \forall\, i \neq j$$

Examples

$$\underline{v}_1 = \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \underline{v}_2 = \begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}$$

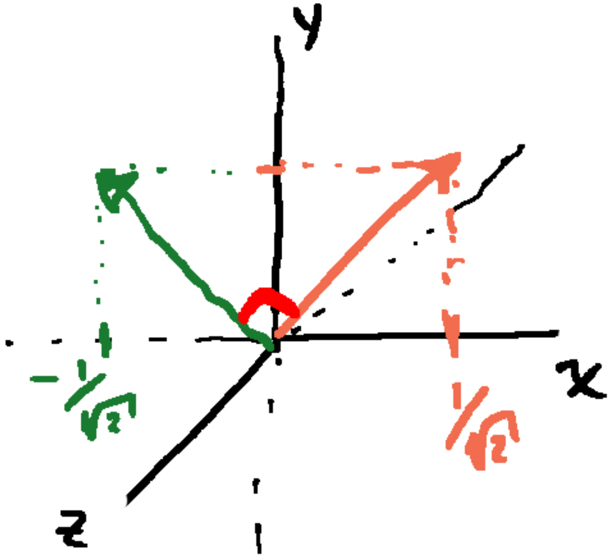


basis for x-y plane

2-D subspace in $\mathbb{R}^3$

$$\underline{f} = \underline{v}_1 w_1 + \underline{v}_2 w_2$$

$$\underline{v}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ 1 \\ 0 \end{bmatrix}, \quad \underline{v}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$



$$\underline{v}_1^T \underline{v}_1 = \underline{v}_2^T \underline{v}_2 = 1$$

$$\underline{v}_1^T \underline{v}_2 = 0$$

orthonormal basis for x-y plane

3

$$\underline{v}_1 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \end{bmatrix}, \quad \underline{v}_2 = \frac{1}{2} \begin{bmatrix} 1 \\ -1 \\ -1 \\ 1 \end{bmatrix}$$

$$\underline{v}_3 = \frac{1}{2} \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix}$$

$$\underline{v}_i^T \underline{v}_i = 1$$

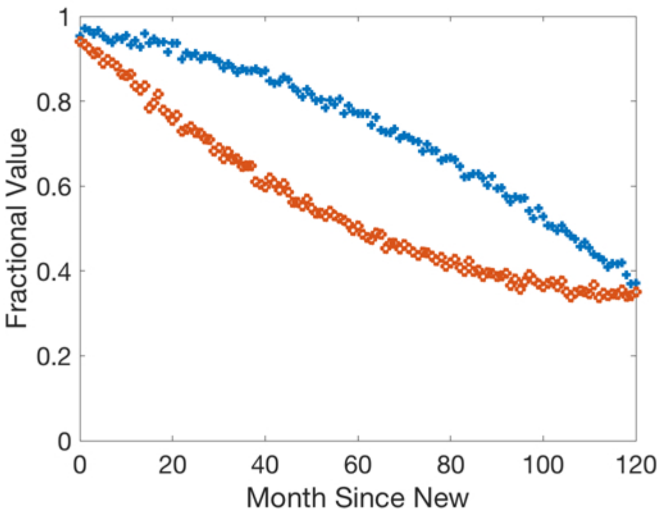
$$\underline{v}_1^T \underline{v}_2 = 0$$

$$\underline{v}_1^T \underline{v}_3 = 0$$

$$\underline{v}_2^T \underline{v}_3 = 0$$

orthonormal basis

Example: Modeling Depreciation



$$p_i = w_1 + t_i w_2 + t_i^2 w_3$$

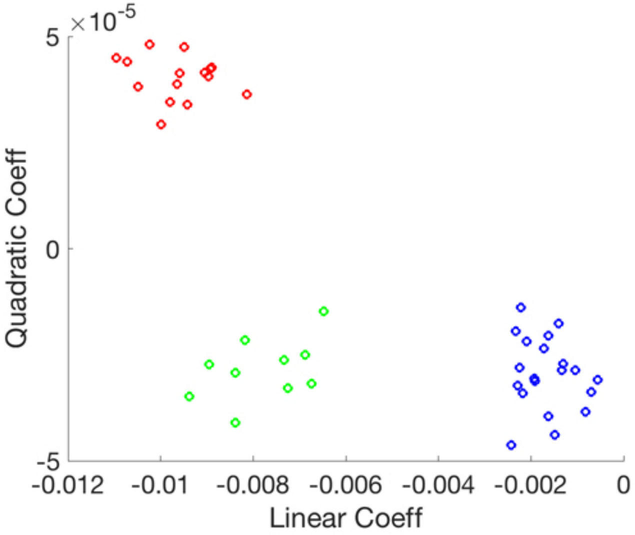
$\underline{p}$: 121 dim

$\underline{w}$: 3 dim

$$\begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_{120} \end{bmatrix} = \begin{array}{c} \begin{bmatrix} 1 & t_0 & t_0^2 \\ 1 & t_1 & t_1^2 \\ \vdots & \vdots & \vdots \\ 1 & t_{120} & t_{120}^2 \end{bmatrix} \\ \begin{array}{ccc} \underline{v}_1 & \underline{v}_2 & \underline{v}_3 \end{array} \end{array} \begin{bmatrix} w_1 \\ w_2 \\ w_3 \end{bmatrix}$$

basis coefficients

$\underline{v}_1$ $\underline{v}_2$ $\underline{v}_3$: bases



loses value initially

loses value

holds value initially

Example: Movie Ratings

$$\begin{bmatrix} \text{last Jedi} \\ \text{Star Trek} \\ \text{When Harry met Sally} \\ \text{Ground Hog Day} \end{bmatrix} \begin{array}{l} \left.\begin{array}{l} \\ \\ \end{array}\right] \text{scifi} \\ \left.\begin{array}{l} \\ \\ \end{array}\right] \text{rom com} \end{array}$$

carrie fisher

Taste profiles (bases)

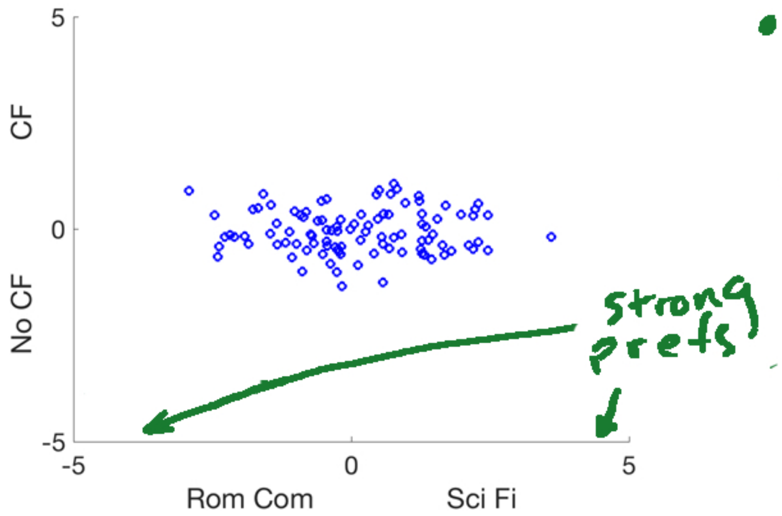
$$\underline{t}_1 = \begin{bmatrix} 1 \\ 1 \\ -1 \\ -1 \end{bmatrix} \qquad \underline{t}_2 = \begin{bmatrix} 1 \\ -1 \\ 1 \\ -1 \end{bmatrix}$$

sf vs. rc

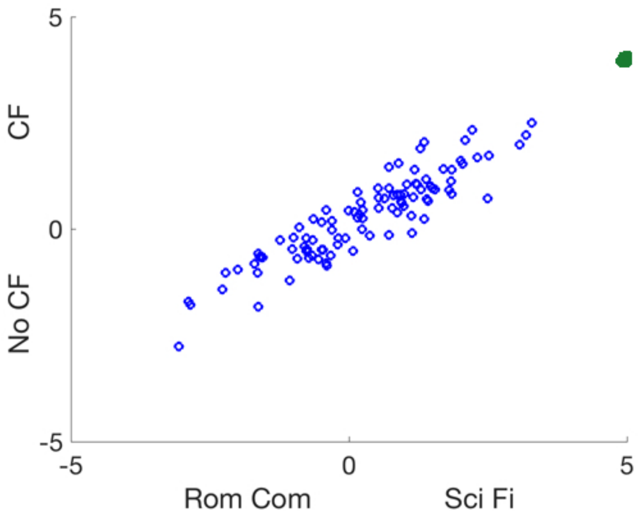
cf vs no cf

$$\underline{r}_i = \underline{t}_1 w_1 + \underline{t}_2 w_2$$

100 people × 4 dimensions



- sf vs rc stronger
- no link



- sf-rc linked with cf-ncf

Open Issues

- Choosing a good/useful basis

- Choosing dimension

- Finding basis coefficients given bases

- Finding orthonormal bases