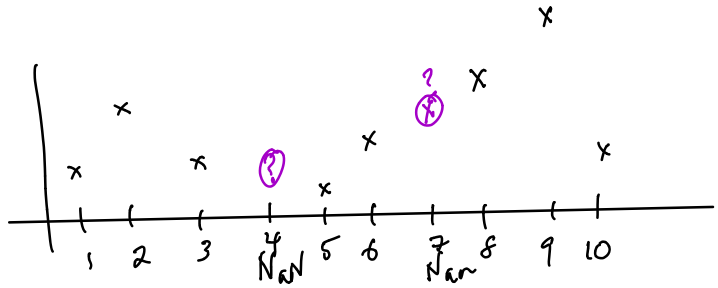
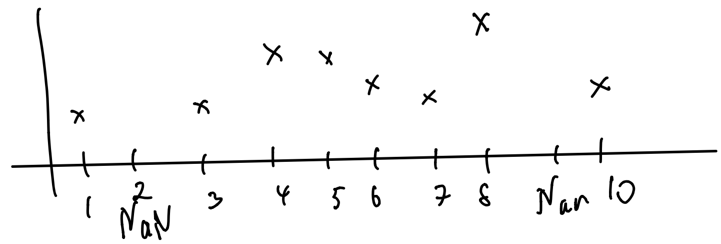
Data set $\{x^{(\mu)}\}_{\mu=1}^{P}$ (NETFLIX PROBLEM)

e.g.) $x^{(1)}$



$x^{(2)}$



Model missing data with the introduction of masks $\{m^{(\mu)}\}_{\mu=1}^{P}$.

For above data we have:

$$m^{(1)} = (1\ 1\ 1\ 0\ 1\ 1\ 0\ 1\ 1\ 1)^T$$
$$m^{(2)} = (1\ 0\ 1\ 1\ 1\ 1\ 1\ 1\ 0\ 1)^T$$

$$m_i^{(\mu)} = \begin{cases} 1 & \text{if } x_i^{(\mu)} \text{ is known} \\ 0 & \text{else} \end{cases}$$

Comment: The value at a masked pt will be ignored by the methodology, not viewed as zero.

Let $\beta = \{u^{(1)}, \ldots, u^{(N)}\}$ be a $\overline{\text{o.n.}}$ basis for the data.

$$x^{(m)} = \sum_{k=1}^{N} \alpha_k^{(m)} u^{(k)}$$

Normally we compute

$$\alpha_k^{(m)} = \langle x^{(m)}, u^{(k)} \rangle.$$

$\|x\|^2 = \langle x, x \rangle$

Another view of where this comes from:

$$E = \| x^{(m)} - \sum_{k=1}^{N'} \alpha_k^{(m)} u^{(k)} \|^2$$

$$= \langle x^{(m)} - \sum_k \alpha_k^{(m)} u^{(k)}, x^{(m)} - \sum_j \alpha_j^{(m)} u^{(j)} \rangle$$

$$= \langle x^{(m)}, x^{(m)} \rangle - 2 \sum_k \alpha_k^{(m)} \langle x^{(m)}, u^{(k)} \rangle$$

$$+ \sum_{jk} \alpha_j^{(m)} \alpha_k^{(m)} \underbrace{\langle u^{(k)}, u^{(j)} \rangle}_{\delta_{jk}}$$

$$E = \|x^{(m)}\|^2 - 2 \sum_k \alpha_k^{(m)} \langle x^{(m)}, u^{(k)} \rangle + \sum_j \left(\alpha_j^{(m)}\right)^2$$

$$\frac{\partial E}{\partial \alpha_m} = 0 - 2 \langle x^{(m)}, u^{(m)} \rangle + 2 \alpha_m^{(m)}$$

$$\alpha_m^{(m)} = \langle x^{(m)}, u^{(m)} \rangle \leftarrow$$

The above result ignores the missing data.

Define a new inner product:

$$\|z\|_m^2 = \langle z, z \rangle_m = \sum_{i=1}^{N} z_i^2 m_i$$

i'th component ignored when $m_i = 0$.

$$\hat{E} = \| x^{(m)} - \sum \hat{\alpha}_k^{(m)} u^{(k)} \|_m^2 \qquad \hat{\alpha} \text{ instead of } \alpha$$

$$= \langle x^{(m)}, x^{(m)} \rangle_m - 2 \sum_k \hat{\alpha}_k^{(m)} \langle x^{(m)}, u^{(k)} \rangle_m$$

$$+ \sum_{jk} \hat{\alpha}_k \hat{\alpha}_j \langle u^{(k)}, u^{(j)} \rangle_m$$

$$\frac{\partial \hat{E}}{\partial \alpha_m} = -2 \langle x^{(m)}, u^{(m)} \rangle_m$$

$$+ \sum_{jk} \left(\delta_{mk} \hat{\alpha}_j^{(m)} + \delta_{mj} \hat{\alpha}_k^{(m)} \right) \langle u^{(k)}, u^{(m)} \rangle_m$$

$$= -2 \langle x^{(m)}, u^{(m)} \rangle_m + \sum_j \hat{\alpha}_j^{(m)} \langle u^{(m)}, u^{(j)} \rangle_m$$

$$+ \sum_k \hat{\alpha}_k^{(m)} \langle u^{(k)}, u^{(m)} \rangle_m$$

$$= -2 \langle x^{(m)}, u^{(m)} \rangle_m + 2 \sum_j \hat{\alpha}_j^{(m)} \langle u^{(m)}, u^{(j)} \rangle_m$$

Since $\partial \hat{E} / \partial \hat{\alpha}_m = 0$ $\boxed{D = ?}$

$$\Rightarrow \sum_{j=1}^{D} \hat{\alpha}_j^{(m)} \langle u^{(m)}, u^{(j)} \rangle_{\underline{m}} = \langle x^{(m)}, u^{(m)} \rangle_{\underline{m}}$$

Define $M_{ij} = \langle u^{(i)}, u^{(j)} \rangle_{\underline{m}}$

$$f_i^{(m)} = \langle x^{(m)}, u^{(i)} \rangle_m$$

$$\hat{\alpha}^{(m)} = (\alpha_1^{(m)}, \ldots, \alpha_D^{(m)})^T$$

So

$$\boxed{M \hat{\alpha}^{(m)} = f^{(m)}}$$

We are solving for $\hat{\alpha}^{(m)}$ w/ gappy data.

$$\begin{cases} x^{(m)} \tilde{=} \sum_{i=1}^{D} \hat{\alpha}_i^{(m)} u^{(i)} \\ \text{Solving } \underset{D\times D}{M} \underset{D\times 1}{\hat{\alpha}^{(m)}} = \underset{D\times 1}{f^{(m)}} \end{cases}$$

In contrast to

$$x^{(n)} \tilde{=} \sum_{i=1}^{D} \alpha_i^{(n)} u^{(i)}$$

$$\alpha_i = \langle x^{(n)}, u^{(i)} \rangle.$$ pays attention to the missing data

E.g. Consider the data set

$$\underline{X} = \begin{bmatrix} 1 & 1 & 2 & -1 \\ 2 & -1 & 1 & 4 \\ 3 & -2 & 1 & 7 \\ 4 & 1 & 5 & 2 \end{bmatrix}$$ rank 2.

Observe the point

$$x = \begin{pmatrix} 6 \\ -3 \\ -7 \\ 9 \end{pmatrix}$$ but get $$\begin{pmatrix} 6 \\ -3 \\ \text{NaN} \\ 9 \end{pmatrix}.$$

$[x = x_1 + 5x_2]$

take $\beta = \{u^{(1)}, \dots, u^{(4)}\}$ SVD

$\sigma_1 = 10.28$, $\sigma_2 = 5.68$
$\sigma_3 = \sigma_4 = 0$.

Don't Use Gappy Approach

$M = I$ since the basis is o.n.

1D reconstruction

$\alpha_1 = -3.4781$

$x \approx [0.11 \quad 1.55 \quad 2.55 \quad 1.78]$

2D reconstruction $\alpha_2 = 10.04$

$x \approx [4.7 \quad -.18 \quad -1.9 \quad 9.3]$

4D $x \approx [6 \quad -3 \quad 0 \quad 9]$

gets gap.

1D

$$M_{11} = .463$$

$$f_1 = -3.48$$

$$\hat{\alpha}_1 = -7.51$$

$$x \approx \begin{bmatrix} 0.25 & 3.35 & 5.5 & 3.85 \end{bmatrix}$$

2D

$$M_{ij} = \langle u^{(i)}, u^{(j)} \rangle_m$$

$$= \begin{bmatrix} 0.46 & -0.32 \\ -.32 & 0.80 \end{bmatrix}$$

$$f = \begin{bmatrix} -3.48 \\ 10.04 \end{bmatrix}$$

Sol. $\hat{\alpha} = \begin{bmatrix} 1.65 \\ 13.12 \end{bmatrix}$

note: that $\hat{\alpha}_1$ is now changed in the 2D case.

2D

$x\,app = [6 \quad -3 \quad -7 \quad 9]$

Summary of "How to Repair Data with a good basis"

- Given an observed pattern x determine the gaps, ie, construct $\underline{m}$.
- identify $\beta = \{u^{(1)}, \dots, u^{(D)}\}$, $\underline{\underline{D}}$.
- Solve $M\alpha = f$ for $D \times D$ problem.
- $g = V_D \hat{\alpha}$ gappy reconstruction.

  solve $M\hat{\alpha} = f$
- repair vector r

$$r = \begin{cases} x_i & m_i = 1 \\ g_i & m_i = 0. \end{cases}$$

Repairing a Set of Gappy Patterns

Observe gappy data matrix

$$Y = \left[y^{(1)} \mid \cdots \mid y^{(p)} \right]$$

associated masks (given or determined)

$$M = \left[m^{(1)} \mid \cdots \mid m^{(p)} \right]$$

NO BASIS!

e.g.]

$$Y = \begin{array}{c} \\ \\ 6 \rightarrow \\ \\ \\ \end{array} \begin{bmatrix} 1 & 2 & 4 & 0 \\ 2 & 3 & 5 & 7 \\ \boxed{0} & 4 & 6 & 8 \\ 4 & 10^{8} & 8 & 0 \\ 5 & 7 & 0 & 17 \end{bmatrix}, \quad M = \begin{bmatrix} 1 & 1 & 1 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ 1 & 1 & 0 & 1 \end{bmatrix}$$

$$\frac{4+6+8}{3}$$

Initial repair:

$$\left(\underline{r}_0^{(\mu)}\right)_i = \begin{cases} y_i^{(\mu)} & m_i^{(\mu)} = 1 \\ \frac{1}{P_i} \sum_{\nu=1}^{p} y_i^{(\nu)} m_i^{(\nu)}, & m_i^{(\mu)} = 0 \end{cases}$$

$$P_i = \sum_{\nu=1}^{p} m_i^{(\nu)}$$

Zeroth Repair:

$$R_0 = \left[\underline{r}_0^{(1)} \mid \underline{r}_0^{(2)} \mid \cdots \mid r_0^{(p)}\right]$$

Approximate Basis U_0:

$$R_0 = U_0 S V_0^T.$$

Now use $\mathcal{U}_0$ to repair R_0.

for each μ.

$$\text{Solve } M \hat{\alpha}_D^{(\mu)} = f^{(\mu)} \quad D \times 1$$

$$g = \mathcal{U}_D \hat{\alpha}_D$$

$$\left(\underline{r}_1^{(\mu)} \right)_i = \begin{cases} y_i & m_i = 1 \\ g_i & m_i = 0. \end{cases}$$

New repair $R_1 = [\underline{r}_1^{(1)} | \dots | r_1^{(P)}]$

Update basis: $\mathcal{U}_1$

$$R_1 = \mathcal{U}_1 \Sigma_1 V_1^T.$$

Summary :

i) make initial repair w/ averaging

ii) Find initial good basis U_0 via svd on R_0

iii) use U_0 to repair R_0 to get R_1

iv) find a new improved basis U_1
where $U_1 S_1 V_1^T = R_1$

repeat.

Stopping criterion at iteration $k+1$

$$\| \underline{\sigma}_k^D - \underline{\sigma}_{k+1}^D \| \leq \varepsilon$$

$$\underline{\sigma}^D = (\sigma_k^D , \ldots , \sigma_K^D).$$