

The problem we are trying to analyse, post-ReLU low rank matrix completion, can be formulated as follows:

The d by n true data matrix M can be given as the product of an orthogonal d by r matrix U and a r by n weight matrix W . We are able to observe M through the mask $\Omega = M \geq 0$. So $\Omega = UW \geq 0$.

So $\Omega_{a,b}$ is defined by $\text{sign}(U_{a,:} \cdot W_{:,b})$, where only the +1s are sampled.

We will assume that W is generated as follows: Each column of W is a random unit vector in R^r , and each column of W is generated independently from other columns of W .

Given that, consider some set S of entries of Ω where S_{1_a} is the first index of the first element of S and S_{1_b} is the second index. So $S_1 = \Omega_{S_{1_a}, S_{1_b}}$.

Let's suppose that all elements in the set have a different second index, so $\forall i, j \in \{1..|S|\}, i \neq j, S_{i_b} \neq S_{j_b}$. Then for a bunch of z , let $z_i = \{-1, 1\}$ arbitrarily.

For some i , $P(S_i = z_0 | \forall S_j \neq S_i, S_j = z_j) = P(S_i = z_0)$ is the independence criteria we're trying to prove: basically, we want the same probability despite knowing all other entries. $z_j = S_j = \text{sign}(U_{S_{j_a},:} \cdot W_{:,S_{j_b}})$. Because z_j is completely defined by this equation, knowing z_j at best reveals $U_{S_{j_a},:}$ and $W_{:,S_{j_b}}$. Because $i_b \neq j_b$ for all j , $W_{:,S_{i_b}}$ is independent from all observed $W_{:,S_{j_b}}$ because of the way in which W was sampled. $W_{:,S_{i_b}}$ is also clearly independent from the $U_{S_{j_a},:}$. So because the z_j are entirely determined by those columns of W and rows of U , and $W_{:,S_{i_b}}$ is independent from those items, $P(W_{:,S_{i_b}} = x | \forall S_j \neq S_i, S_j = z_j) = P(W_{:,S_{i_b}} = x)$; that is, the column of W is independent from all observations made which use different columns of W .

$$\begin{aligned} & P(S_i = z_0 | \forall S_j \neq S_i, S_j = z_j) \\ &= P(\text{sign}(U_{S_{i_a},:} \cdot W_{:,S_{i_b}}) = z_0 | \forall S_j \neq S_i, S_j = z_j) \\ &= \sum_{V \in Gr(r,d)} P(U = V | \forall S_j \neq S_i, S_j = z_j) P(\text{sign}(U_{S_{i_a},:} \cdot W_{:,S_{i_b}}) = z_0 | U = V, \forall S_j \neq S_i, S_j = z_j) \\ &= \sum_{V \in Gr(r,d)} P(U = V | \forall S_j \neq S_i, S_j = z_j) * \frac{1}{2} \text{ because the probability that an independent uniform random unit vector lies on one side of a fixed half plane is always } \frac{1}{2} \\ &= \frac{1}{2} \text{ because we are summing over the entire Grassmanian distribution. So any set in the mask where each element has a unique second index is independent.} \end{aligned}$$