Gọi là các user, là các item, và là rating của user cho item , (: rating matrix).

Phương pháp **user-based** nearest neighbor recommendation

Độ tương tự của hai user u và v được tính bởi công thức sau:

Dự đoán rating của user u cho item i được tính như sau:

Phương pháp **item-based** nearest neighbor recommendation

Độ tương tự của hai item i và j được tính bởi công thức sau

Dự đoán rating của user u cho item i được tính như sau:

Trong đó

* , là rating trung bình của user u và v
* là tập những user tương tự với user u
* là tập những rating của user u

Cho rating matrix sau:

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| u\i | i1 | i2 | i3 | i4 | i5 |
| u | 5 | 3 | 4 | 4 | ? |
| u1 | 3 | 1 | 2 | 3 | 3 |
| u2 | 4 | 3 | 4 | 3 | 5 |
| u3 | 3 | 3 | 1 | 5 | 4 |
| u4 | 1 | 5 | 5 | 2 | 1 |

a. Tính theo phương pháp user-based nearest neighbor (dùng 1 nearest neighbor).

b. Tính theo phương pháp item-based nearest neighbor (dùng 1 nearest neighbor).

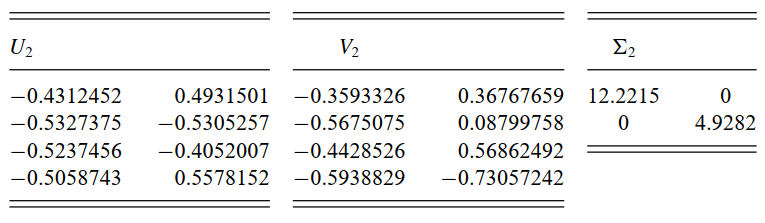
SVD-based recommendation

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| i\u | u1 | u2 | u3 | u4 |
| i1 | 3 | 4 | 3 | 1 |
| i2 | 1 | 3 | 2 | 6 |
| i3 | 2 | 4 | 1 | 5 |
| i4 | 3 | 3 | 5 | 2 |

Phương pháp SVD dựa trên việc một ma trận M bất kỳ có thể tách thành

* U và V được gọi là left và right singular vectors
* là ma trận đường chéo, được gọi singular values.

Cho ma trận rating như trên, áp dụng SVD, ta tính và chỉ giữ lại k đặc trưng quan trọng nhất (thường k << m, và k << n). Khi k = 2, ta được kết quả như hình bên dưới.



Cho biết rating vector (5, 3, 4, 4) của một user u sẽ trở thành như thế nào trong không gian mới (2 chiều)?