Matrix Completion

Objectives

- define the matrix completion problem
- approach missing data using low-rank models
- introduce iterative singular value thresholding

Use "patterns" to fill in missing entries

Ratings matrix

$\underline{X} \in \mathbb{R}^{N \times M}$

$$
\begin{bmatrix}
5 & 4 & 9 & 1 & \times \\
9 & 6 & \times & \times & 7 \\
\times & 10 & \times & 2 & 4 \\
3 & 7 & 3 & \times & \times \\
8 & \times & \times & 6 & 2
\end{bmatrix}
$$

N movies ↓

M users →

We need to infer them from the existing data (predict)

Can we predict the missing entries?

★ Model: assume $\underline{X}$ is well approximated with a small number of patterns

$$\underline{X} \approx \sum_{i=1}^{r} \underline{t}_i \underline{s}_i^T = \underline{T}\,\underline{S}$$

Rank-r matrix

small number of factors affecting the ratings

genres, actors, director...

hobbies, age, address...

Matrix completion: use known data to find patterns and predict missing entries

$\Omega = \{ i,j : \underline{X}_{ij} \text{ given} \}$ indices of known values

(M需要和X中的已知值一样, rest of them should lead to min rank)

☆(low-rank 说明用户的喜好可以被少数几个"基本模式"解释)

1) Rank minimization $\underline{X} = \underset{\underline{M}}{\operatorname{argmin}} \operatorname{rank}(\underline{M})$ s.t. $\underline{M}_{ij} = \underline{X}_{ij} \ \forall i,j \in \Omega$

Minimizing rank = Minimizing sum of patterns

minimum number of patterns matching given values

Intractable!

$\operatorname{rank} \underline{M} = \# \{ \ell : \sigma_\ell > 0 \}$

= number of positive singular values

2) Nuclear norm minimization

$\underline{X} = \underset{\underline{M}}{\operatorname{argmin}} \|\underline{M}\|_* \ \underset{\text{subject to}}{\underbrace{\text{s.t.}}} \ \underline{M}_{ij} = \underline{X}_{ij} \ \forall i,j \in \Omega$

Nuclear/trace norm

$\|\underline{M}\|_* = \sum_\ell \sigma_\ell$

(Sum of singular values is the nuclear norm)

And we want to minimize the sum

Computationally tractable

Recommended

Iterative Singular Value Thresholding is one possible algorithm

Initialize

$\underline{M}^{(0)} = \underline{0}$ Initialize M to all-zero

Set threshold or r

some threshold

num of rank

Iterate

for $k = 1, 2, 3, \dots$

$\underline{M}^{(k)} = \underline{M}^{(k-1)}$

$\underline{M}^{(k)}_{\Omega} = \underline{X}_{\Omega}$ (fill in known values)

$[\underline{U}, \underline{\Sigma}, \underline{V}] = \mathrm{svd}(\underline{M}^{(k)})$

$$\hat{\Sigma}_{ii} = \Sigma_{ii} \cdot \begin{cases} 1 & \Sigma_{ii} > \text{threshold} \\ 0 & \Sigma_{ii} \le \text{threshold} \end{cases}$$

强模式变成1

弱模式变成0

—or—

$$\hat{\Sigma}_{ii} = \begin{cases} \Sigma_{ii}, & i \le r \\ 0, & i \ge r+1 \end{cases}$$

第k轮的最佳猜测

$\underline{M}^{(k)} = \underline{U}\,\underline{\hat{\Sigma}}\,\underline{V}^T$

清理过的新Σ

if $\| \underline{M}^{(k)} - \underline{M}^{(k-1)} \|_F < \varepsilon$

stop

else

next k

把k轮的$M^{(k)}$和上一轮的$M^{(k-1)}$相比，如果变化很小说明converge停止，否则继续

Matrix completion is an open problem

- choosing r or threshold in ISVT
- multiple algorithms:
  - convergence
  - complexity
  - noise
- results depend on distribution of missing entries
- applications include missing pixels in images, position from partial distance info, ...


DILBERT
SCOTT ADAMS
dilbert.com
WE WON A GOVERN-MENT CONTRACT TO MEASURE OCEAN TEMPERATURES.
WHICH PART OF THE OCEAN?
THE WHOLE OCEAN.
WE CAN'T PUT SENSORS EVERY-WHERE IN THE OCEAN. IT'S TOO BIG.
WE CAN MEASURE A BUNCH OF PLACES AND ESTIMATE THE REST.
SO . . . YOU WANT ME TO MEASURE 1% OF THE OCEAN'S TEMPERATURE AND ESTIMATE THE OTHER 99%?
Twitter: @scottadamssays
I DON'T KNOW HOW TO DO THAT.
TRY USING MATH.
© 2019 Scott Adams, Inc./Dist. by Andrews McMeel
WOULDN'T IT BE CHEAPER TO MEASURE NOTHING AND JUST ESTIMATE THE WHOLE THING?
3-10-19
Dilbert.com
EVERY NOW AND THEN YOU COME UP WITH A GREAT IDEA.