Q4: How Does Netflix Recommend Movies?

Mung Chiang

Netflix

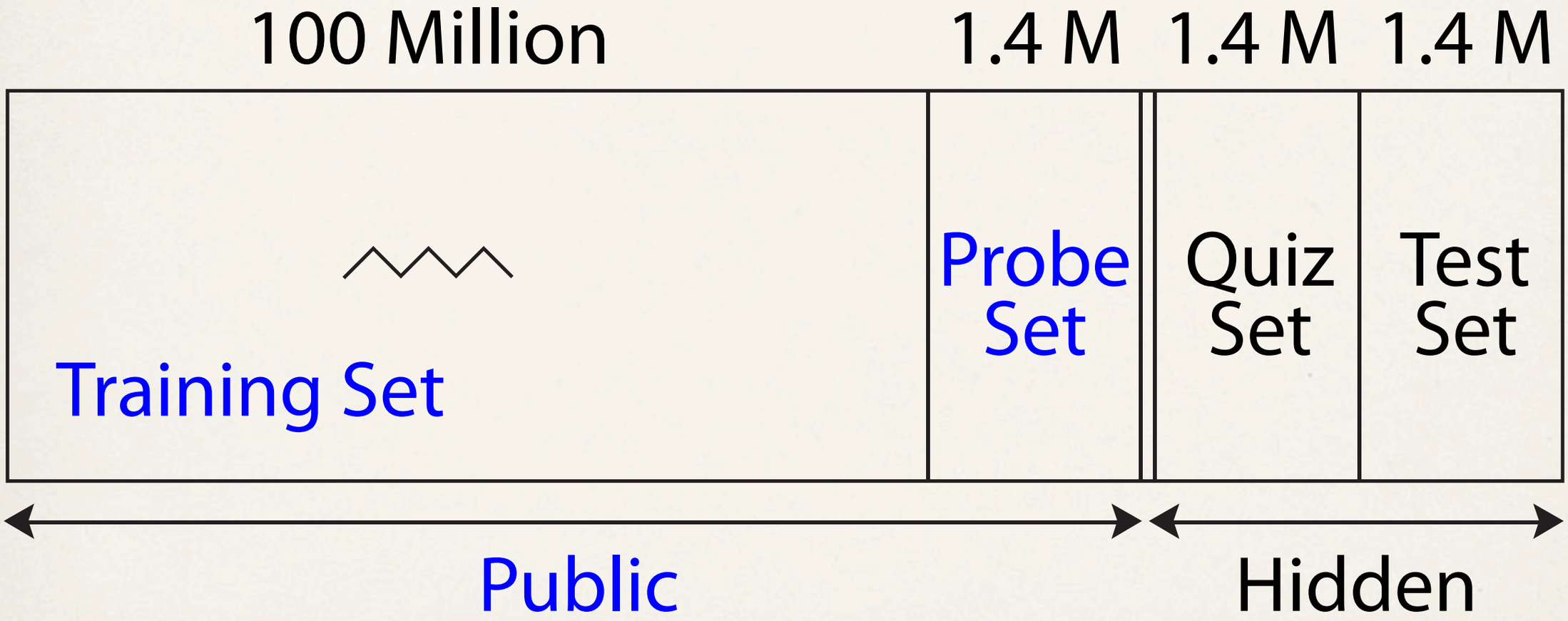
- DVD rental
- Online streaming
  - 30% of Internet traffic in summer 2011
- 23 million subscribers
- Recommendation accuracy important
  - Compare to Amazon, YouTube, Pandora

Recommendation Problem

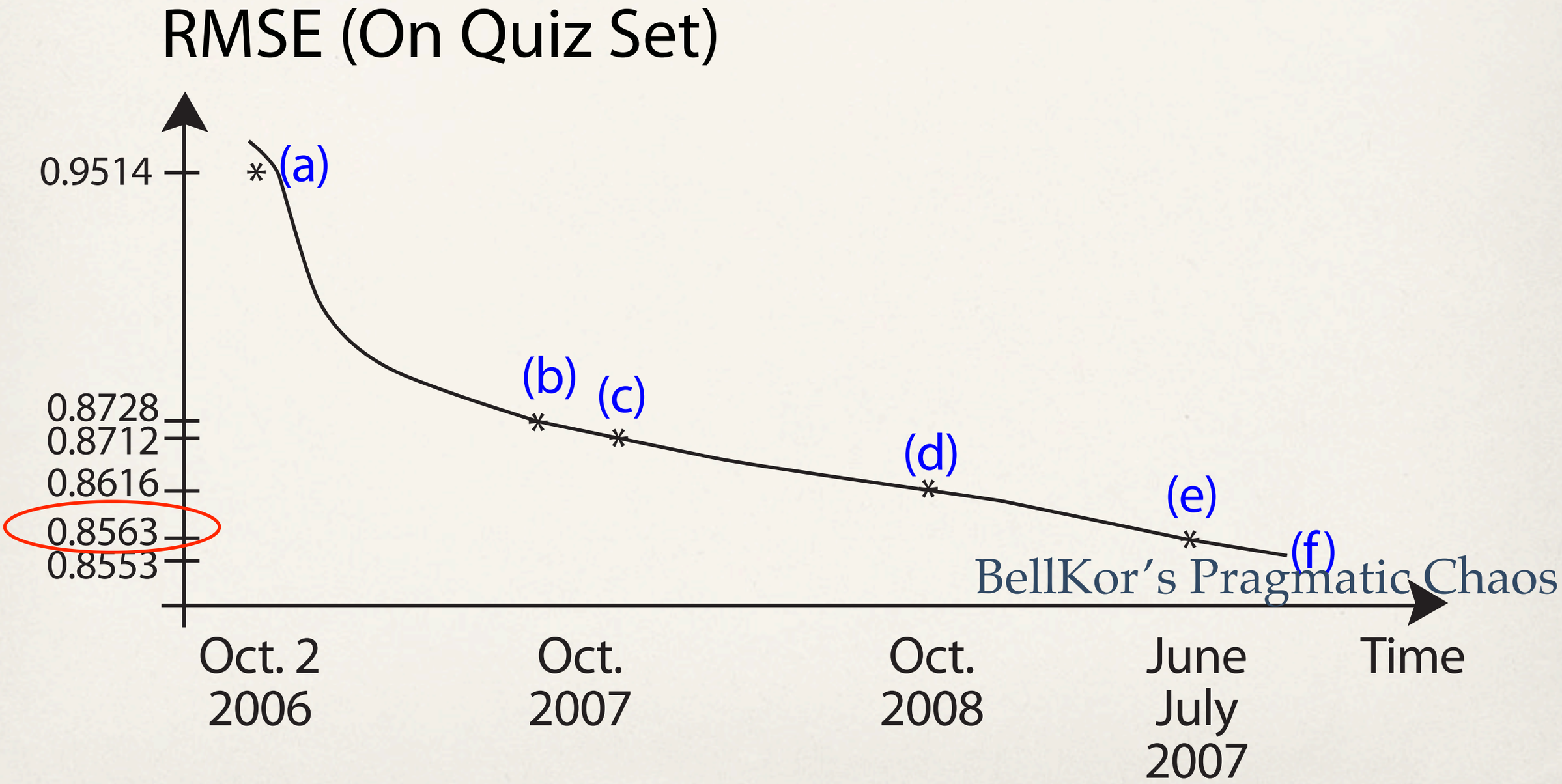
- Input
  - [User, Movie, 1-5 stars, Time]
- Output
  - Prediction of ratings
- Metric
  - RMSE

$$\sqrt{\sum_{(u,i)} \frac{(r_{ui} - \hat{r}_{ui})^2}{C}}$$

Netflix Prize



Brief History



Matrix, or Weighted Bipartite

Movies

Users	1	2	3	4	5	6	7	8
1		5		2	4			
2	4		3	1			3	
3		5	4		5		4	
4						1	1	2
5	3		?		?	3		
6		?	2		4		?	

Challenges

- Size
  - 480,000 users, 17,770 movies, 100 million ratings
- Sparsity
  - Only 1% dense, most users do not rate most movies
- Diversity
  - Most users rate only few movies, few users rate many movies

Key Approaches

- Content based filtering (local)
- Collaborative filtering (global)
  - Neighborhood method (statistical correlation)
    - Movie-Movie, or User-User similarities
  - Latent factor method (structural simplicity)
- Implicit feedback, temporal dynamics...

Baseline Predictor

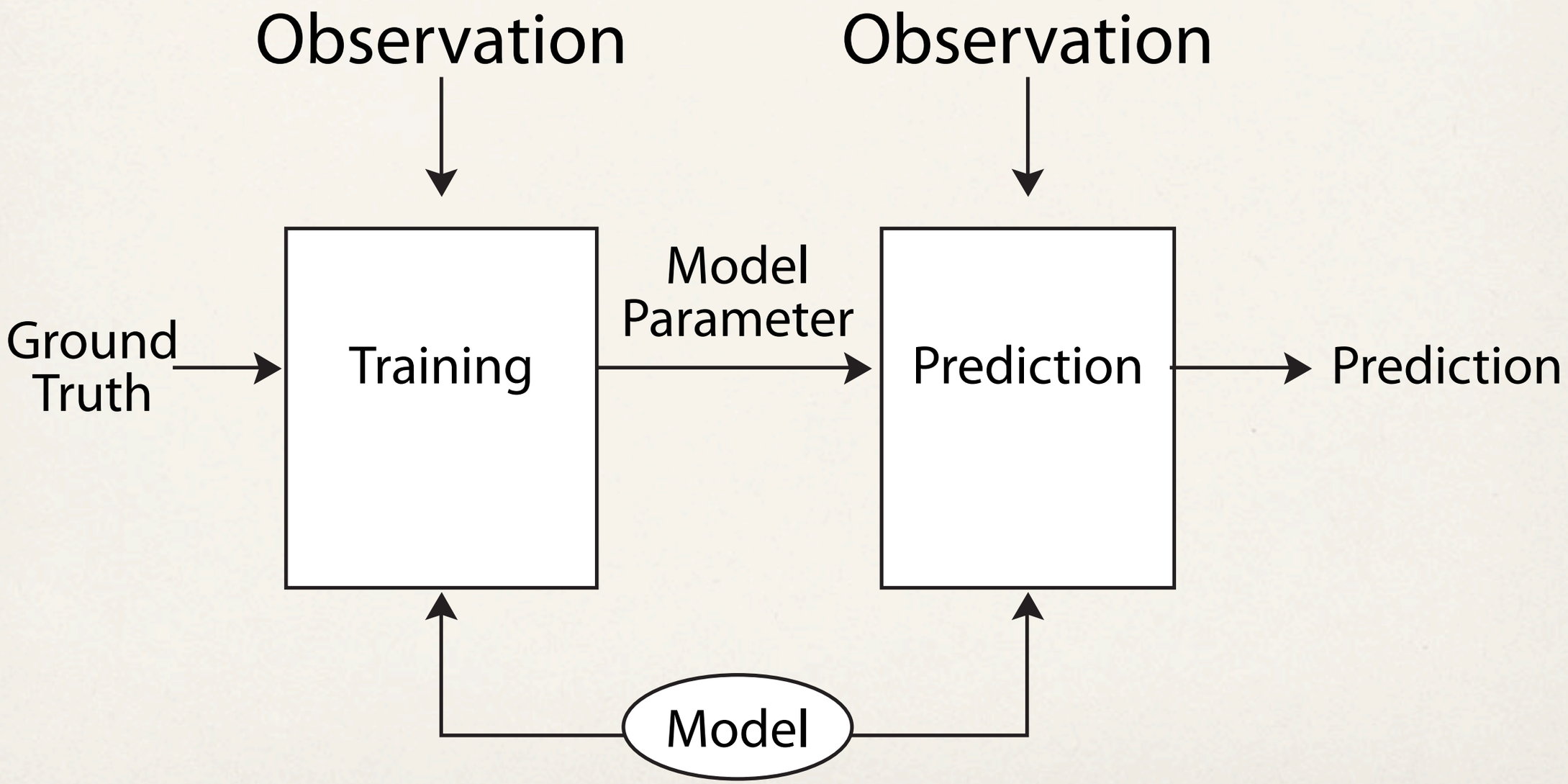
Average

$$\hat{r}_{ui} = \bar{r} + b_u + b_i$$

User bias

Movie bias

Training -> Prediction



Parameter Training

Training Data

$$\underset{\{b_u, b_i\}}{\text{minimize}} \sum_{(u,i)} \left(r_{ui} - \hat{r}_{ui}\right)^2$$

Model Parameters

Model Prediction

Least Squares -> Linear Eq.

Variables

Minimize $\|\mathbf{Ab} - \mathbf{c}\|_2^2$

Optimal choice

$$(\mathbf{A}^T\mathbf{A})\mathbf{b} = \mathbf{A}^T\mathbf{c}$$

Similarity Metric

$$\tilde{r}_{ui} = r_{ui} - \hat{r}_{ui} = r_{ui} - (\bar{r} + b_u + b_i)$$

$$d_{ij} = \frac{\sum_u \tilde{r}_{ui}\tilde{r}_{uj}}{\sqrt{\sum_u (\tilde{r}_{ui})^2 \sum_u (\tilde{r}_{uj})^2}}$$

Neighborhood Method

Neighborhood predictor

Baseline predictor

$$\hat{r}_{ui} = (\bar{r} + b_u + b_i) + \frac{\sum_{j \in \mathcal{L}} d_{ij} \tilde{r}_{uj}}{\sum_{j \in \mathcal{L}} |d_{ij}|}$$

Fix a (user, movie) pair

Set of top L neighbors

Example: Input

$$\mathbf{R} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{pmatrix} 5 & 4 & 4 & - & \mathbf{5} \\ - & 3 & 5 & \mathbf{3} & 4 \\ 5 & 2 & - & \mathbf{2} & 3 \\ - & \mathbf{2} & 3 & 1 & 2 \\ 4 & - & \mathbf{5} & 4 & 5 \\ \mathbf{5} & 3 & - & 3 & 5 \\ 3 & \mathbf{2} & 3 & 2 & - \\ 5 & \mathbf{3} & 4 & - & 5 \\ \mathbf{4} & 2 & 5 & 4 & - \\ \mathbf{5} & - & 5 & 3 & 4 \end{pmatrix} \end{array}$$

Least Squares Solution

$$
\begin{array}{c} \\ 1 \\ 2 \\ \vdots \\ 30 \end{array}
\begin{array}{c}
\begin{array}{cccccccc} 1 & 2 & \cdots & 10 & A & B & \cdots & E \end{array} \\
\begin{pmatrix}
1 & 0 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
0 & 1 & \cdots & 0 & 1 & 0 & \cdots & 0 \\
\vdots & & & & \ddots & & & \vdots \\
0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 1
\end{pmatrix}
\end{array}
\begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_{10} \\ b_A \\ b_B \\ \vdots \\ b_E \end{pmatrix}
-
\begin{pmatrix} r_{1A} - \bar{r} \\ r_{2A} - \bar{r} \\ \vdots \\ r_{10E} - \bar{r} \end{pmatrix}
$$

$$\mathbf{b}_u^* = [0.62, 0.42, -0.28, -1.78, 0.52, 0.49, -1.24, 0.45, 0.40, 0.23]^T$$

$$\mathbf{b}_i^* = [0.72, -1.20, 0.60, -0.60, 0.33]^T$$

Baseline Predictor

$$\hat{\mathbf{R}} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{pmatrix} 5.00 & 3.09 & 4.90 & - & \mathbf{4.62} \\ - & 2.89 & 4.69 & \mathbf{3.49} & 4.42 \\ 4.10 & 2.19 & - & \mathbf{2.78} & 3.71 \\ - & \mathbf{1.00} & 2.49 & 1.29 & 2.22 \\ 4.90 & - & \mathbf{4.79} & 3.58 & 4.51 \\ \mathbf{4.88} & 2.96 & - & 3.56 & 4.48 \\ 3.15 & \mathbf{1.23} & 3.03 & 1.82 & - \\ 4.84 & \mathbf{2.92} & 4.72 & - & 4.44 \\ \mathbf{4.84} & 2.92 & 4.72 & 3.51 & - \\ \mathbf{4.61} & - & 4.49 & 3.29 & 4.22 \end{pmatrix} \end{array}$$

Shift by Baseline Predictor

$$\tilde{\mathbf{R}} = \mathbf{R} - \hat{\mathbf{R}} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{pmatrix} 0 & 0.91 & -0.90 & - & ? \\ - & 0.11 & 0.31 & ? & -0.42 \\ 0.90 & -0.19 & - & ? & -0.71 \\ - & ? & 0.51 & -0.29 & -0.22 \\ -0.90 & - & ? & 0.42 & 0.49 \\ ? & 0.040 & - & -0.56 & 0.52 \\ -0.15 & ? & -0.031 & 0.18 & - \\ 0.16 & ? & -0.72 & - & 0.56 \\ ? & -0.87 & 0.33 & 0.54 & - \\ ? & - & 0.51 & -0.29 & -0.22 \end{pmatrix} \end{array}$$

Similarity

$$
\begin{aligned}
d_{BC} &= \frac{\tilde{r}_{1B}\tilde{r}_{1C} + \tilde{r}_{2B}\tilde{r}_{2C} + \tilde{r}_{9B}\tilde{r}_{9C}}{\sqrt{(\tilde{r}_{1B}^2 + \tilde{r}_{2B}^2 + \tilde{r}_{9B}^2)(\tilde{r}_{1C}^2 + \tilde{r}_{2C}^2 + \tilde{r}_{9C}^2)}} \\
&= \frac{(0.91)(-0.90) + (-0.11)(0.31) + (-0.87)(0.33)}{\sqrt{(0.91^2 + 0.11^2 + 0.87^2)(0.90^2 + 0.31^2 + 0.33^2)}} \\
&= -0.84.
\end{aligned}
$$

$$
\mathbf{D} = \begin{array}{c} \\ A \\ B \\ C \\ D \\ E \end{array}
\begin{array}{c}
\begin{array}{ccccc} A & B & C & D & E \end{array} \\
\begin{pmatrix}
- & -0.21 & -0.41 & -0.97 & -0.75 \\
-0.21 & - & -0.84 & -0.73 & 0.51 \\
-0.41 & -0.84 & - & -0.22 & -0.93 \\
-0.97 & -0.73 & -0.22 & - & 0.68 \\
-0.75 & 0.51 & -0.93 & 0.68 & -
\end{pmatrix}
\end{array}
$$

Neighborhood Prediction

$$\hat{r}_{3D} = (\bar{r}+b_3+b_D)+\frac{d_{DA}\tilde{r}_{3A}+d_{DB}\tilde{r}_{3B}}{|d_{DA}|+|d_{DB}|} = 2.78+\frac{-0.97*0.90+(-0.73)*(-0.19)}{0.97+0.73} = 2.35.$$

$$\hat{\mathbf{R}}_{neighborhood} = \begin{array}{c} \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{array} \begin{array}{c} \begin{array}{ccccc} A & B & C & D & E \end{array} \\ \begin{pmatrix} 5.00 & 3.99 & 3.99 & - & \mathbf{5.00} \\ - & 2.58 & 4.86 & \mathbf{3.38} & 4.11 \\ 4.81 & 2.19 & - & \mathbf{2.35} & 2.81 \\ - & \mathbf{1.00} & 2.71 & 1.29 & 1.71 \\ 4.46 & - & \mathbf{4.30} & 4.49 & 5.42 \\ \mathbf{4.97} & 3.52 & - & 3.52 & 4.48 \\ 2.97 & \mathbf{1.16} & 3.03 & 1.97 & - \\ 4.28 & \mathbf{3.64} & 4.16 & - & 4.77 \\ \mathbf{4.25} & 2.44 & 5.54 & 4.33 & - \\ \mathbf{4.87} & - & 4.71 & 3.29 & 3.71 \end{pmatrix} \end{array}$$

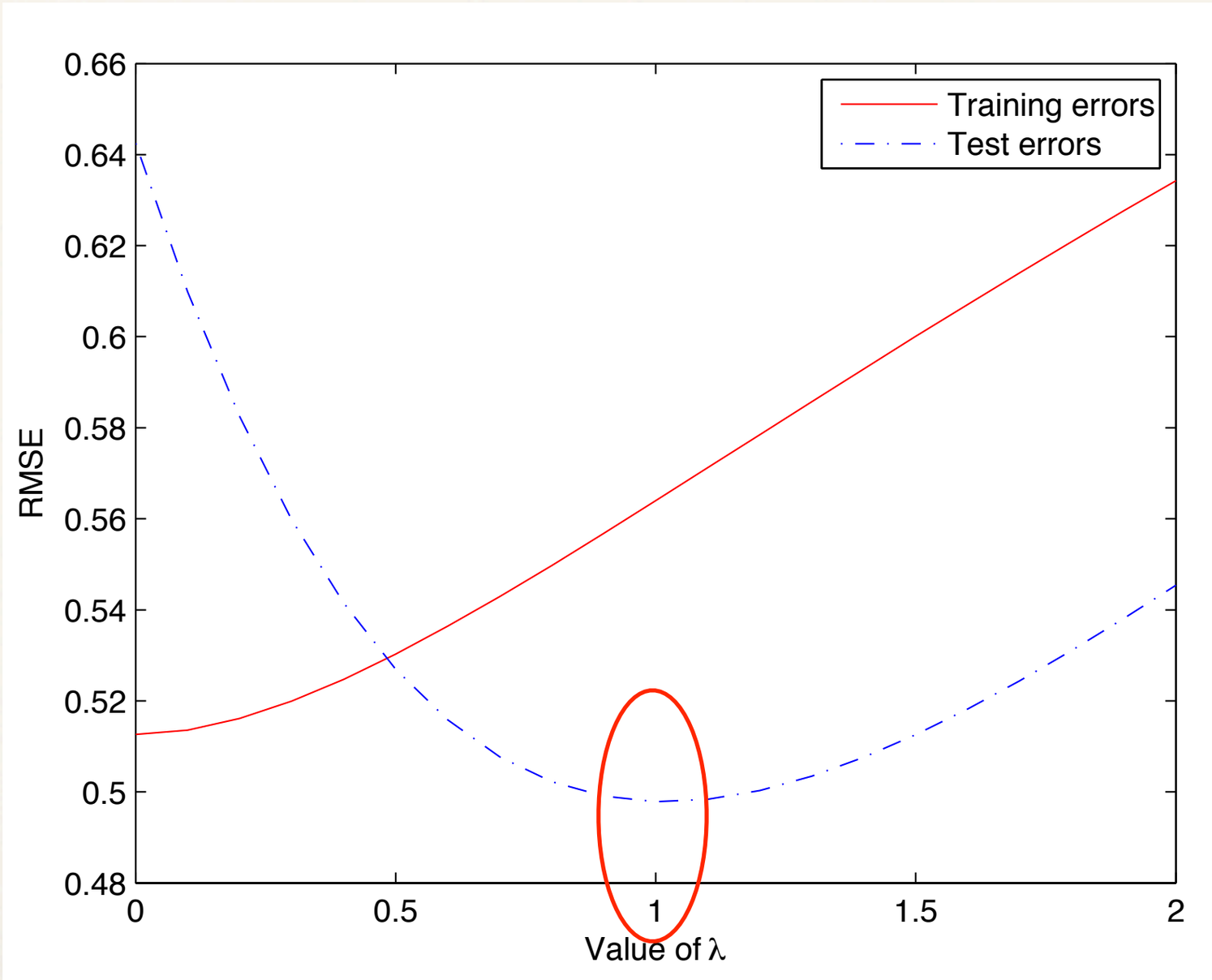
RMSE on Quiz Set: drop from 0.64 to 0.54

Bonus: Regularization

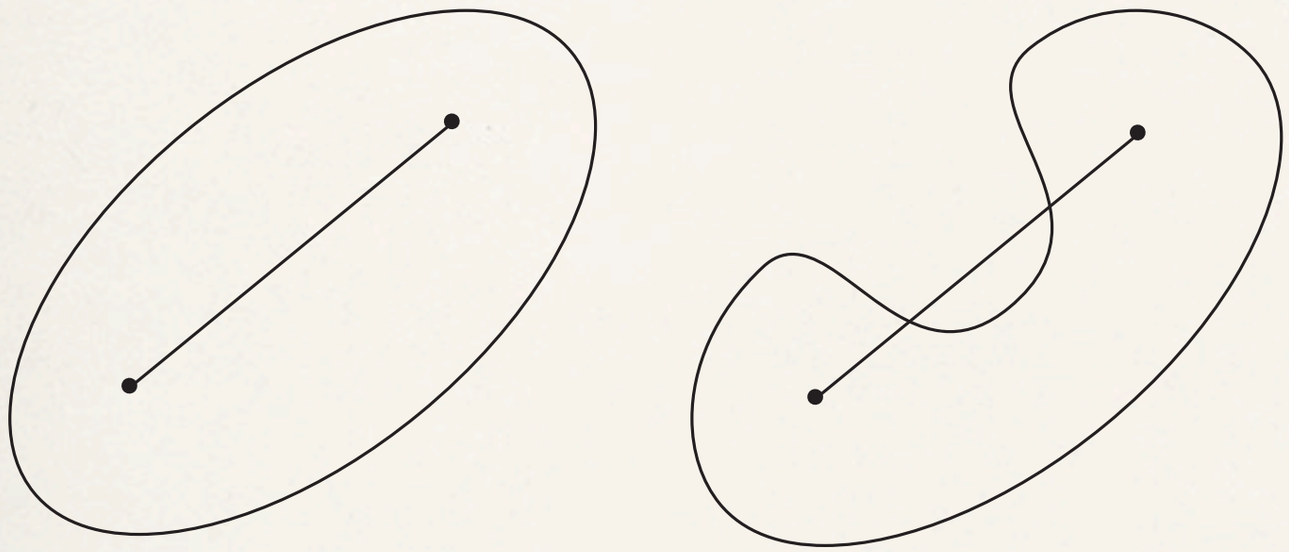
Regularization Weight

$$\text{minimize}_{\{b_u, b_i\}} \sum_{(u,i)} (r_{ui} - \hat{r}_{ui})^2 + \lambda \left(\sum_u b_u^2 + \sum_i b_i^2 \right)$$

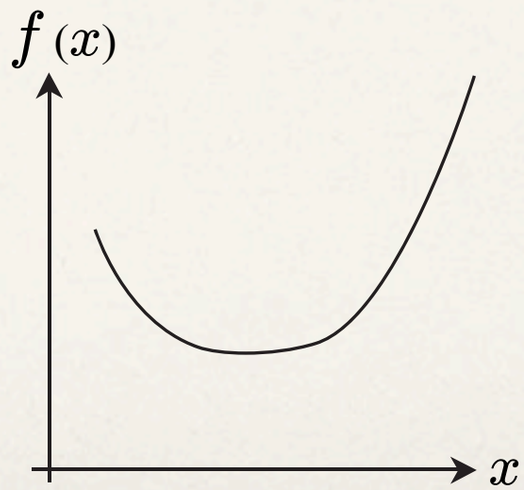
Regularization



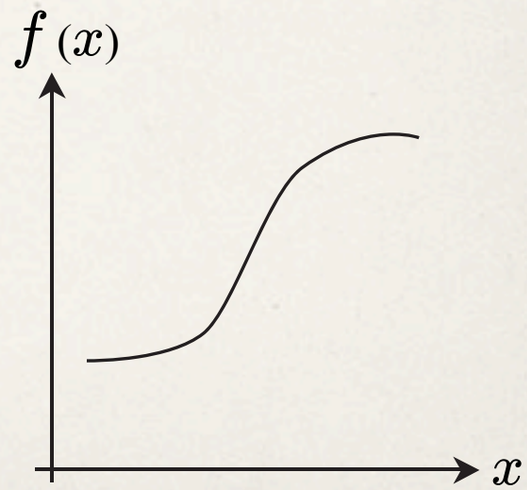
Bonus: Convex Optimization



(a) (b)



(a)



(b)

Key Phrase Summary

- Collaborative filtering
  - Neighborhood method
    - Similarity metric
- Least squares
  - Regularization
- Convex optimization