

# Sistemas de recomendacoes

Alex Grilo

March 8, 2015

## Introducao

O objetivo deste relatorio e descrever os assuntos estudados durante a disciplina de MC032 no segundo semestre de 2010. O principal interesse deste trabalho foi pesquisar na literatura sobre o problema de sistemas de recomendacoes e diferentes abordagens para resolve-lo. Na secao 2 e descrito em linhas gerais o que e um sistema de recomendacoes. Nas secoes 3 e 4 sao mostradas as principais ideias encontradas nos trabalhos, respectivamente. Na secao 5 encontra-se a conclusao do estudo e ideias para trabalhos futuros.

## Sistemas de recomendacoes

Um sistema de recomendacoes tem como objetivo recomendar a usuarios produtos que o agradem. Para isso, o sistema de recomendacoes se baseia inicialmente em respostas ja conhecidas de outros usuarios e alguma informacao sobre o usuario que deseja obter uma recomendacao. O conjunto de informacao inicial que o sistema de recomendacoes ja possui e utilizado para comparar com a informacao obtida do usuario em questao para tentar deduzir quais outros produtos interessam ao usuario.

## Competitive Recommendation Systems

### Introducao

A ideia basica desta abordagem e reduzir o problema de sistema de recomendacoes no problema de reconstrucao de matrizes a partir de informacoes parciais da mesma. Foram utilizadas estrategias de reconstrucao de matrizes baseadas na tecnica de SVD, encontrando boas aproximacoes para a matriz original

*Divisao em tipos* Assume-se que os usuarios podem ser divididos em grupos com interesses semelhantes, classificando-os em tipos. Consideraremos que  $k$  seja o numero de tipos, onde  $k$  e uma constante que nao depende do numero de usuarios.

### Qualidade dos algoritmos

Um bom algoritmo de recomendacoes deve dar boas recomendacoes utilizando o minimo de informacao possivel. Sao estabelecidas 2 metricas para medir a qualidade de um algoritmo de recomendacoes.

Um algoritmo e *c – competitivo para amostragem* se usa somente  $ck$  linhas e colunas da matriz para recomendar mais produtos.

Um algoritmo e *f – competitivo para utilidade* se faz boas recomendacoes a  $fm$  usuarios, onde  $m$  e o numero de usuarios.

### Notacao

Nesta secao, a seguinte notacao sera utilizada:

A matriz de recomendacao original

$A_{ij}$  valor da utilidade do produto  $j$  para o usuario  $i$

$A(i)$   $i$ -esima linha da matriz  $A$ , que corresponde ao vetor de utilidades do  $i$ -esimo usuario.

$A(i)$   $i$ -esima coluna a matriz  $A$ , que corresponde a utilidade de um produto para todos os usuarios.

$A_k$  Melhor aproximacao com posto  $k$  de  $A$ , obtida atraves da tecnica de SVD.  
air valor da utilidade do  $r$ -esimo produto com maior utilidade para o usuario  $i$

#### *Reconstrucao de matrizes*

##### *Escolha das linhas*

A escolha de um numero constante de linhas segue a ideia de que sao necessarios alguns usuarios respondendo um questionario sobre todos os produtos para que os outros usuarios sejam classificados a partir das respostas destes. Na pratica, empresas pagam pessoas para o preenchimento desses formularios.

Portanto a partir da escolha das  $ak$  linhas, podemos formar uma matriz  $V$  de tipos efetivos.

##### *Escolha das colunas*

A ideia de escolher um numero constante de colunas acontece pois todo usuario que deseja obter uma recomendacao deve responder inicialmente um pequeno questionario, dando alguma informacao sua para que o algoritmo de recomendacao

#### **Ideia da prova**

Primeiramente separa-se os casos em que o usuario faz parte do comite e quando o usuario nao faz Para o caso em que o usuario faz parte do comite, ele tera que fazer  $K$  provas. Para o caso em que o usuario nao esta no comite, e calculada a esperanca do numero de provas que ele tera que fazer. Como a ordem dos produtos recomendados e aleatoria, o numero de provas que um usuario tera que fazer e uma variavel aleatoria geometrica. Logo, basta fazer a analise da esperanca para essa variavel aleatoria. O resultado do teorema aparece com a esperanca do numero de provas para cada caso.

*Comite* A tecnica de criar um comite do algoritmo proposto possui serias desvantagens. A primeira delas e que a complexidade de recomendacao individual e  $\Omega(n)$ , ja que os usuarios do comite tem que provar todos os algoritmos. Um segundo ponto contra esta tecnica e a existencia de usuarios desonestos, que, quando escolhidos para o comite, atrapalham o funcionamento do algoritmo de recomendacoes.

#### **Conclusao e trabalhos futuros**

Atraves deste estudo, verificou-se a existencia de solucoes com abordagens totalmente diferentes para o problema de sistemas de recomendacoes. Mas o mais importante foi o aprendizado das tecnicas mais gerais para resolucao de problemas como por exemplo SVD. Potenciais trabalhos futuros sao a implementacao e comparacao dos algoritmos na pratica ou alteracao neles para que ao inves de dispor de informacoes de colunas inteiras, os algoritmos funcionem com posicoes esparsas na matriz.