CSE 8803RS: Recommendation Systems

Lecture 13: Joint Low-Rank Matrix Factorizations

Hongyuan Zha

School of Computational Science & Engineering
College of Computing
Georgia Institute of Technology

Recommender Systems for Documents

Google scholar | matrix computations | Search | Advanced Scholar Search | Scholar Preferences

Scholar | Articles and patents | anytime | include citations | Results **1** - **10** of about **1,280,000**. (**0.15** sec)

[BOOK] **Matrix computations**
GH Golub, CF Van Loan - 1996 - books.google.com
©1983, 1989, 1996 The Johns Hopkins University Press All rights reserved. Published 1996 Printed in the United States of America on acid-free paper 9876 First edition 1983 Second edition 1989 Third edition 1996 The Johns Hopkins University Press 2715 North Charles Street **...**
Cited by 29117 - Related articles - Library Search - All 12 versions

[PDF] SPARSKIT: A basic tool kit for sparse **matrix computation** — psu.edu [PDF]
Y Saad - 1994 - Citeseer
Abstract. This paper presents the main features of a tool package for manipulating and working with sparse matrices. One of the goals of the package is to provide basic tools to facilitate exchange of software and data between researchers in sparse **matrix computations**. Our starting **...**
Cited by 576 - Related articles - View as HTML - All 18 versions

[BOOK] Fundamentals of **matrix computations**
DS Watkins - 2002 - books.google.com
This text is printed on acid-free paper. © Copyright © 2002 by John Wiley & Sons, Inc., New York. All rights reserved. Published simultaneously in Canada. No part of this publication may be reproduced, stored in a retrieval system or transmitted in any form or by any means, **...**
Cited by 542 - Related articles - All 15 versions

[CITATION] **Matrix computations**. Johns Hopkins Studies in the Mathematical …
GH Golub, CF Van Loan - Johns Hopkins University Press, 1996
Cited by 445 - Related articles

Recommender Systems for Documents

Google scholar [Search] Advanced Scholar Search
Scholar Preferences

Scholar Results **1** - **10** of about **101** related to **Golub: Matrix computations**. (**0.10** sec)

[BOOK] Matrix computations
GH Golub, CF Van Loan - 1996 - books.google.com
©1983, 1989, 1996 The Johns Hopkins University Press All rights reserved. Published 1996 Printed in the United States of America on acid-free paper 9876 First edition 1983 Second edition 1989 Third edition 1996 The Johns Hopkins University Press 2715 North Charles Street ...
Cited by 29117 - Related articles - Library Search - All 12 versions

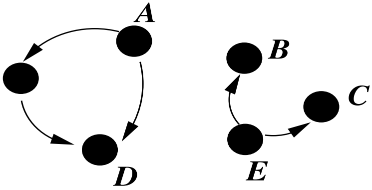
[BOOK] The algebraic eigenvalue problem
JH Wilkinson - 1988 - books.google.com
NUMERICAL MATHEMATICS AND SCIENTIFIC COMPUTATION *P. Dierckx: Curve and surface fittings with splines *H. Wilkinson: The algebraic eigenvalue problem *I. Duff, A. Erisman, and J. Reid: Direct methods for sparse matrices *JD Pryce: Numerical solution of ...
Cited by 5961 - Related articles - Library Search - All 10 versions

[BOOK] The symmetric eigenvalue problem
BN Parlett - 1998 - books.google.com
SIAM's Classics in Applied Mathematics series consists of books that were previously allowed to go out of print. These books are republished by SIAM as a professional service because they continue to be important resources for mathematical scientists. Editor-in-Chief Robert E. O' ...
Cited by 2570 - Related articles - All 11 versions

[BOOK] LAPACK Users' guide
E Anderson, Z Bai, C Bischof, S Blackford, J Demmel, J ... - 1999 - books.google.com
7/lie 5etie5 inclucles li2siclbOo!<5 2lic^ 5aftwsse guicles 25 well 25 mcoioßtsplis on plActica! implementÄtion c>f computatioiiZ! metnc>6z, environments, Znci tc»c>>z. I'lie focus >5 c»n M2><inß tecent Developments 2V2i!Z,ole in 2 ps2ctic2> fosMÄt tc» seseÄrcliesz ...
Cited by 3604 - Related articles - Library Search - All 8 versions

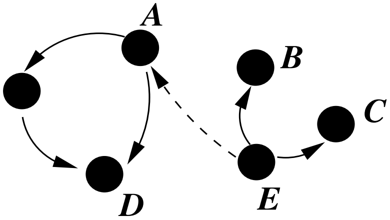
[BOOK] Matrix analysis
RA Horn, CR Johnson - 1990 - books.google.com
Contents Preface Chapter 0 Review and miscellanea 0.0 Introduction 0.1 Vector spaces 0.2 Matrices 0.3 Determinants 0.4 Rank 0.5 Nonsingularity 0.6 The usual inner product 0.7 Partitioned matrices 0.8 Determinants again 0.9 Special types of matrices 0.10 Change of basis ...
Cited by 10430 - Related articles - Library Search - All 5 versions

Recommender Systems Based on Citation Graphs



- Citation graph:
  — documents ⇒ vertices
  — citations ⇒ directed edges
- Document similarity based on co-citations:
  — B and C are similar because they are both cited by E
- Citation graph is *sparse* and *noisy*

Issues Using a Single Citation Graph



(a) Missing citations (b) Same authors

Exploring Multiple Data Sources

- Real-world problems are complex and involves data from multiple sources
- Consider three relationships:
  — Citation relationship: a directed graph
  — Author-Document relationship: a bipartite graph
  — Document-Venue relationship: a bipartite graph
- A special case of Entity-Relationship Model (ERM)
  — Entity types: documents, authors, and venues
  — Relationships: citation, authorship, document-venue

Simultaneous Profiling of Entities from Multiple Sources

- Profiling: assigning a set of numeric/categorical features to entities
- Embedding: projecting entities to k-dimensional Euclidean space
  — metric properties reflect semantics
- Using, e.g., nearest neighbor search, we can
  — for a given document, which are the "closest" documents:
  Google Scholar: *related document* function
  — for a given document, who are the "closest" authors
  — for a given author, who are the "closest" documents
  — . . .
- Document classification, clustering and visualization

Relationships and Profiling

Relationships as Graphs:

- Citation relationship: $G_D = (V_D, E_D)$
- Author-Document relationship: G_{AD} or its adjacency matrix
- Document-Venue relationship: G_{VD} or its adjacency matrix

Entity Profiles:

- Documents: $F_D \in R^{n_d \times k}$, n_d number of documents
- Authors: $F_A \in R^{n_a \times k}$, n_a number of authors
- Venues: $F_V \in R^{n_v \times k}$, n_v number of venues

Optimization Problem for Simultaneous Profiling

Objective function:

$$L(F_D, F_A, F_V) = \sum_{(i,j)\in E_D} \|F_D(i,:) - F_D(j,:)\|^2 + \|G_{AD} - F_A F_D^T\|_F^2$$

$$+\|G_{VD} - F_V F_D^T\|_F^2 + \underbrace{\lambda_A\|F_A\|_F^2 + \lambda_D\|F_D\|_F^2 + \lambda_V\|F_V\|_F^2}_{\text{regularization terms}}$$

Optimization Problem:

$$\min_{F_D,F_A,F_V} \quad L(F_D, F_A, F_V)$$

Regularized SVD for Multiple Matrices

- With three types of entities U, V, W, and three relations A, B, C, the objective function is

$$E(U,V,W) = \frac{1}{2}\sum_{(i,j)\in O_A}(A_{ij}-\sum_{k=1}^{K}U_{ik}V_{jk})^2+\frac{1}{2}\sum_{(i,j)\in O_B}(B_{ij}-\sum_{k=1}^{K}V_{ik}W_{jk})^2$$

$$+\frac{1}{2}\sum_{(i,j)\in O_C}(C_{ij}-\sum_{k=1}^{K}W_{ik}U_{jk})^2+\frac{\tilde{\lambda}}{2}\sum_{i,k}U_{ik}^2+\frac{\tilde{\lambda}}{2}\sum_{i,k}V_{ik}^2+\frac{\tilde{\lambda}}{2}\sum_{i,k}W_{ik}^2$$

- O_A, O_B, O_C index sets of observed entries for A, B, C
- The objective function is in additive form. The gradient can again be computed one component at a time, and then sum the results

For Each Iteration

For each (i, j) $\in O_A$

Compute the current estimate $\hat{A}_{ij} = u_i v_j^T$

Compute the current error $R_{ij} = A_{ij} - \hat{A}_{ij}$

For each k = 1, ..., K

$$U_{ik} \leftarrow U_{ik} + \mu(R_{ij} V_{jk} - \lambda U_{ik})$$
$$V_{jk} \leftarrow V_{jk} + \mu(R_{ij} U_{ik} - \lambda V_{jk})$$

For each (i, j) $\in O_B$

Compute the current estimate $\hat{B}_{ij} = v_i w_j^T$

Compute the current error $R_{ij} = B_{ij} - \hat{B}_{ij}$

For each k = 1, ..., K

$$V_{ik} \leftarrow V_{ik} + \mu(R_{ij} W_{jk} - \lambda V_{ik})$$
$$W_{jk} \leftarrow W_{jk} + \mu(R_{ij} V_{ik} - \lambda W_{jk})$$

Algorithm: Pseudo-Code

For Each Iteration

For each (i, j) $\in O_C$

Compute the current estimate $\hat{C}_{ij} = w_i u_j^T$

Compute the current error $R_{ij} = C_{ij} - \hat{C}_{ij}$

For each k = 1, ..., K

$$W_{ik} \leftarrow W_{ik} + \mu(R_{ij} U_{jk} - \lambda W_{ik})$$
$$U_{jk} \leftarrow U_{jk} + \mu(R_{ij} W_{ik} - \lambda U_{jk})$$

Algorithm: Multiple Matrices

- Given A, B, C

$$A \approx UV^T, \quad B \approx VW^T, \quad C \approx WU^T,$$

- Consider the following matrix,

$$\begin{bmatrix} U \\ V \\ W \end{bmatrix} [U^T, V^T, W^T] = \begin{bmatrix} UU^T & UV^T & UW^T \\ & VV^T & VW^T \\ & & WW^T \end{bmatrix}$$

$$\approx \begin{bmatrix} ? & A & C^T \\ & ? & B \\ & & ? \end{bmatrix} = \mathcal{A}$$

Experiments

- Data from Citeseer and DBLP
- Evaluations:
  — Randomly remove documents from citations, and predict missing citations
  — Use F_1 measure: $F_1 = pr/(p+r)$, where p is precision and r is recall
- Use label propagation to rank documents

F-score Comparisons

	f \ m	m=t	m=5	m=10
$DS1$	f(lap)	0.013	0.048	0.192
	f(svd)	0.035	0.086	0.138
	f(new)	**0.108**	**0.242**	**0.325**
$DS2$	f(lap)	0.011	0.046	0.156
	f(svd)	0.027	0.072	0.109
	f(new)	**0.083**	**0.158**	**0.229**

Table 1: The f-score calculated on different numbers of top documents, m.

F-score Comparisons

	f \ t	t=1	t=2	t=3	t=4
$DS1$	f(lap)	0.041	0.048	0.075	0.086
	f(svd)	0.062	0.088	0.099	0.103
	f(new)	**0.197**	**0.242**	**0.248**	**0.252**
$DS2$	f(lap)	0.037	0.047	0.068	0.077
	f(svd)	0.049	0.072	0.082	0.086
	f(new)	**0.121**	**0.158**	**0.181**	**0.182**

Table 2: The f-score w.r.t. different numbers of left-out documents, t.

Low-Rank Factorization with Side Information

Consider A as the interaction matrix between two types of entities. There are possible extra information about the entities

- Extra similarity information:

$$\min_{F,G} \|A - FG^T\|_F^2 + +\lambda(\sum_{ij} w_{ij}^F \|F_i - F_j\|^2 + \sum_{ij} w_{ij}^G \|G_i - G_j\|^2)$$

- Extra features S_F, S_G describing the entities:

$$\min_{F,G} \|A - S_F F G^T S_G^T\|_F^2 + \lambda(\|F\|^2 + \|G\|^2)$$

or

$$\min_{F,G,X_F,X_G} \|A - FG^T\|_F^2 + \lambda(\|S_F - FX_F\|_F^2 + \|S_G - GX_G\|_F^2)$$

General Case: Graph Case

- A collection of different types of entities $X_1, \ldots, X_n$
- A graph $G = (V, E)$
  — $V = \{X_1, \ldots, X_n\}$
  — $(i, j) \in E$ if A_{ij} a relationship between X_i and X_j
- Profiles for each entity: F_i for X_i
- Objective function

$$\sum_{(i,j) \in E} \|A_{ij} - F_i F_j^T\|_F^2 + \text{regularization term}$$

General Case: Hyper-graph Case

- A collection of different types of entities $X_1, \ldots, X_n$
- A collection of relationships among the entities $A_{J_1}, \ldots, A_{J_m}$, where $J_i \subset \{1, 2, \ldots, n\}$
- Profiles for each entity: F_i for X_i
- Objective function

$$\sum_{i=1}^{m} \|A_{J_i} - \sum_{j=1}^{k} F_{s_1}(:, i) \otimes F_{s_2}(:, i) \otimes \cdots \otimes F_{s_{n_i}}(:, i)\|_F^2 +$$

$$+\text{regularization term}$$

Here $J_i = \{s_1, s_2, \ldots, s_{n_i}\}$

- Connections with probabilistic graphical models