

Suvrit Sra

Email: suvrit.sra@tuebingen.mpg.de
Url: <http://www.cs.utexas.edu/~suvrit/>
(O): +49.7071.601.572

Citizen of: India Ph.D. title: **Matrix Nearness Problems in Data Mining**.
Committee: Inderjit Dhillon, Todd Arbogast, Joydeep Ghosh,
 Raymond Mooney, and Vijaya Ramachandran.

Objectives: Develop, analyze and implement efficient algorithms for problems and applications in data mining, machine learning, biology and medicine. To bring optimization theory and machine learning closer to each other.

Education

Ph.D., Computer Science, The University of Texas at Austin. August 2000–August 2007
M.S., Computer Science, The University of Texas at Austin. August 2006
B. Engg. (Hons.), Computer Science, BITS, Pilani, India. July 1995–June 1999

Relevant Work Experience

- Postdoctoral Research Scientist (Bernhard Schölkopf's research group), Max-Planck Institute for Biological Cybernetics: *Oct. 2007–now*
- Intern at Microsoft Research (John Platt's group), Redmond: *(Jun 2005–Aug 2005)*.
- Research Assistant for Dr. I. S. Dhillon during several semesters: *(Spring 2002–now.)*.
- Teaching Assistant for **CS303E** (Introductory Java Programming), **CS336** (Analysis of Programs), **CS383C** (Graduate level Numerical Linear Algebra), and **CS313K** (Introduction to Logic): *(Fall 2001–now)*.
- Software Engineer at Hughes Software Systems, India: *(July 1999–May 2000)*.
- Intern at Tektronix Engg. Dev. Bangalore, India: *(Jan 1999–June 1999)*.

Publications

Note: For all my publications before May 2006, the author names are ordered alphabetically.

- [1] J. Brickell, I. S. Dhillon, S. Sra, and J. A. Tropp. The Metric Nearness Problem. *SIAM J. Matrix Analysis and Appl.*, 2007. *Accepted*.
- [2] D. Kim, S. Sra, and I. S. Dhillon. Fast Projection-Based Methods for the Least Squares Nonnegative Matrix Approximation Problem. *Statistical Analysis and Data Mining*, 1, 2007. *Invited Paper*.
- [3] D. Kim, S. Sra, and I. S. Dhillon. Fast Newton-type Methods for the Least Squares Nonnegative Matrix Approximation Problem. In *SIAM Data Mining*, 2007. Recognized within *Best of SDM* papers.
- [4] J. V. Davis, B. Kulis, P. Jain, S. Sra, and I. S. Dhillon. Information-theoretic Metric Learning. In *ICML*, June 2007. *Best Student Paper*.
- [5] J. V. Davis, B. Kulis, S. Sra, and I. S. Dhillon. Information-theoretic Metric Learning. In *NIPS Workshop on Learning to Compare Examples*, December 2006.
- [6] A. Surendran and S. Sra. Incremental Aspect Models for Mining Document Streams. In *PKDD*, pages 633–640, September 2006.
- [7] S. Sra. Efficient large scale linear programming support vector machines. In *ECML*, pages 767–774, September 2006.
- [8] S. Sra and J. A. Tropp. Row-action methods for compressed sensing. In *Int. Conf. on Acoustics, Speech, and Signal Processing (ICASSP)*. IEEE, May 2006.

- [9] A. Banerjee, I. S. Dhillon, J. Ghosh, and S. Sra. Clustering on the Unit Hypersphere using von Mises-Fisher Distributions. *JMLR*, 6:1345–1382, Sep 2005.
- [10] I. S. Dhillon and S. Sra. Generalized Nonnegative Matrix Approximations with Bregman Divergences. In *Neural Information Processing Systems (NIPS)*, Vancouver, Canada, 2005.
- [11] I. S. Dhillon, S. Sra, and J. A. Tropp. Triangle Fixing Algorithms for the Metric Nearness Problem. In *Neural Information Processing Systems (NIPS)*, 2004.
- [12] H. Cho, I. S. Dhillon, Y. Guan, and S. Sra. Minimum Sum Squared Residue based Co-clustering of Gene Expression data. In *Proc. 4th SIAM International Conference on Data Mining (SDM)*, pages 114–125, Florida, 2004. SIAM.
- [13] A. Banerjee, I. S. Dhillon, J. Ghosh, and S. Sra. Generative model-based clustering of directional data. In *Proceedings of The Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining(KDD-2003)*, pages 19–28, 2003.

Selected Technical Reports

- [14] D. Kim, S. Sra, and I. S. Dhillon. A New Projected Quasi-Newton Approach for Nonnegative Least Squares Problem. Technical Report TR-06-54, Department of Computer Sciences, University of Texas, 2006.
- [15] S. Sra, P. Jain, and I. S. Dhillon. Modeling data using directional distributions: Part II. Technical Report TR-07-05, Department of Computer Sciences, University of Texas, February 2007.
- [16] S. Sra and I. S. Dhillon. Decomposition Based Convex Optimization Methods for Matrix Approximations. Technical report, Department of Computer Sciences, University of Texas, 2006. Under preparation.
- [17] S. Sra and I. S. Dhillon. Nonnegative matrix approximation: Algorithms and applications. Technical Report TR-06-27, Computer Sciences, University of Texas at Austin, 2006.
- [18] I. S. Dhillon, S. Sra, and J. A. Tropp. The Metric Nearness Problems with Applications. Technical Report TR-03-23, Computer Sciences, University of Texas at Austin, 2003.
- [19] I. S. Dhillon and S. Sra. Modeling data using directional distributions. Technical Report TR-03-06, Department of Computer Sciences, University of Texas, January 2003.

Dissertation

S. Sra. **Matrix Nearness Problems in Data Mining**. The Univ. of Texas at Austin (Comp. Sci.). Aug 2007.

Academic Honors and Awards

Best Student Paper, ICML 2007	June 2007
Best of SDM Papers	April 2007
Travel award for SDM 2007	April 2007
Invited talk at SAGE days 3 (at IPAM, UCLA).	February 2007
Topic: <i>Sparse Matrix Computations with SAGE</i>	
Filed two patents for work done at MS Research	2005–2006
Travel grants for NIPS 2005	Dec 2005
Travel grant for SDM 2004	April 2004
Recipient of Microelectronics and Computer Development (MCD) Fellowship	Sep 2000–Aug 2004
Received ≥ 7 checks from Donald Knuth for reporting Errors in his books	July 1999–Jan 2001
100% merit scholarship for being in Top 10 out of 700	Aug 1995–Dec 1996
First rank out of approximately 10,000 candidates (12th grade)	May 1995
Best student in 10th grade	1992

Professional Service

Reviewer: JMLR, NIPS, SIGKDD, SDM, ICML, ICDM, CIKM, SIREV, SIMAX, SISC, NSF & NASA grant proposals.

Software that I have written

SSVD	Driver and wrapper library around SVDPACKC (on netlib)
NNMA	Optimized implementations of Nonnegative Matrix Approximation
SSLIB	Sparse Matrix Manipulation Library
MYEIGS	Eigenvalues and eigenvector library (wrapper for ARPACK)
FSOLVER	Optimization software for Large-scale Linear and Quadratic Programs
BLITZ ^{SVM}	Efficient large-scale linear SVMs (Classification & Regression)
Misc.	See webpage for details on other Software

Computer Skills

Languages:	C/C++, Java, Lisp, Perl, Python, etc.
Environments:	Unix/Linux, Windows, Cygwin, etc.
Tools:	Matlab, Mathematica, CPLEX, MOSEK, GSL, Netbeans, Visual Studio, LAPACK, ARPACK, etc.

Relevant Coursework

Algorithm Analysis and Data Structures, Large Scale Data Mining, Artificial Intelligence, Numerical Linear Algebra, Operating Systems, Programming Languages and Compilers, Advanced Computer Architecture, Discrete Mathematics, Advanced Operating Systems, Machine Learning, Real Analysis, Parallel Algorithms, Pattern Recognition.

Language Abilities

English, Hindi, German, Punjabi; Italian (Beginner)

Additional Academic Interests

(i) Matrix Analysis, Linear Algebra, and Algebra (ii) Convex analysis and optimization (iii) Non-linear Optimization (iv) Asymptotic analysis, approximating integrals (v) Discrete mathematics, enumerative combinatorics (vi) Programming

Other Interests

(i) Hiking, Weight Training, Swimming, (ii) Health and Nutrition, (iii) German language and Culture, (iv) Learning languages, (v) Music, Reading, Writing, (vi) Philosophy, Psychology.

References (5)

Inderjit S. Dhillon

inderjit@cs.utexas.edu

Associate Professor

Dept. of Computer Sciences

The University of Texas at Austin

Austin, Texas, 78712. USA

Arun C. Surendran

acsuren@microsoft.com

Senior Applied Researcher

Microsoft adCenter Labs

Microsoft Corporation

Redmond, WA, 98052. USA

Joydeep Ghosh

ghosh@ece.utexas.edu

Professor

Dept. of Electronics and Computer Engg.

The University of Texas at Austin

Austin, Texas, 78712. USA

Arindam Banerjee

banerjee@cs.umn.edu

Assistant Professor

Dept. of Computer Science & Engg.

University of Minnesota, Twin Cities

Minneapolis, MN 55455. USA

Vijaya Ramachandran

vlr@cs.utexas.edu

Professor

Dept. of Computer Sciences

The University of Texas at Austin

Austin, Texas, 78712. USA