CURRICULUM VITAE

Updated on November 11, 2009

Hao Helen Zhang
Department of Statistics
North Carolina State University
Campus Box 8203
Raleigh, NC 27695-8203

Phone: 919-513-4859
Fax: 919-515-1169
Email: hzhang@stat.ncsu.edu
URL: www4.stat.ncsu.edu/∼hzhang

Education/Training

University of Wisconsin - Madison	Statistics	Ph.D. 2002. Adviser: Grace Wahba
Vanderbilt University	Mathematics	Graduate study, 1996-1997.
Peking University	Mathematics	B.Sc. 1996

Positions and Employment

Aug 2008-present	Associate Professor	Department of Statistics, North Carolina State University
Aug 2002-July 2008	Assistant Professor	Department of Statistics, North Carolina State University
2003-2004	Faculty Fellow	Statistical and Applied Mathematical Sciences Institute, NC.
1998-2002	Research Assistant	Department of Statistics, University of Wisconsin-Madison
1997-1998	Teaching Assistant	Department of Statistics, University of Wisconsin-Madison
1996-1997	Teaching Assistant	Department of Mathematics, Vanderbilt University

Editorial Service

Jan 2008-present	Associate Editor of *Journal of American Statistical Association*
July 2008-present	Associate Editor of *Biometrics*
Sep 2009-present	Associate Editor of *Statistical Analysis and Data Mining*

Other Experience and Professional Memberships

2001-present	Member, American Statistical Association (ASA)
2001-present	Lifetime Member, Institute of Mathematical Statistics (IMS)
2002-present	Member, International Biometric Society, ENAR
1997-2001	Member, American Mathematical Society (AMS)

Honors and Awards

2007	National Science Foundation Faculty Early Career Development (CAREER) Award
2004	Invited to Mathematisches Forschungsinstitut Oberwolfach, Germany
2003	Institute of Mathematical Statistics (IMS) Laha Travel Award
2003	Faculty Research and Professional Development Award, North Carolina State University
2000-2001	Vilas Professional Development Fellowship, University of Wisconsin-Madison
1996-1997	University Fellowship, Vanderbilt University
1994-1995	Guanghua Award for Outstanding Undergraduate, Peking University

Refereed Journal Publications

(available from http://www4.stat.ncsu.edu/∼hzhang/publications)

1. Wu, Y., **Zhang, H. H.**, and Liu, Y. (2009) Robust Model-free Multiclass Probability Estimation. *Journal of American Statistical Association*, to appear.

2. Qiao, X., **Zhang, H. H.**, Liu, Y., Todd, M. and Marron, S. (2009) Weighted distance weighted discrimination and its asymptotic properties. *Journal of American Statistical Association*, to appear.

3. Storlie, C., Bondell, H., Reich, B. and **Zhang, H. H.** (2009) The adaptive COSSO for nonparametric surface estimation and model selection. *Statistica Sinica*, to appear.

4. Xiao, N., Zhang, D., and **Zhang H. H.** (2009) Variable Selection for Semiparametric Mixed Models in Longitudinal Studies, *Biometrics*, to appear.

5. Hwang, W. Y., **Zhang, H. H.**, and Ghosal, S. (2009) FIRST: Combining forward iterative selection and shrinkage in high dimensional sparse linear regression. *Statistics and Its Interface*, **2**, 341-348.

6. Xiao, N., **Zhang, H. H.**, and Zhang D. (2009) Automatic Model Selection for Partially Linear Models, *Journal of Multivariate Analysis*, **100**, 2100-2111.

7. Zou, H. and **Zhang, H. H.** (2009) On the adaptive elastic-net with a diverging number of parameters. *Annals of Statistics*, **37**, 1733-1751.

8. **Zhang, H. H.** (2008) Discussion of "Sure independence screening for ultra-high dimensional feature space", *Journal of the Royal Statistical Society, Series B*, **70**.

9. Liu, H., Tang, Y. and **Zhang, H. H.** (2008) A new chi-square approximation to the distribution of non-negative definite quadratic form in non-central normal variables. *Computational Statistics and Data Analysis*, **53**, 853-856.

10. **Zhang, H. H.** (2008) Support Vector Machine Classification for High Dimensional Microarray Data Analysis, with Applications in Cancer Research. In *High-Dimensional Data Analysis in Cancer Research*, Li and Xu, eds. Springer-Verlag.

11. **Zhang, H. H.**, Liu, Y., Wu, Y. and Zhu, J. (2008) Variable selection for multicategory SVM via sup-norm regularization. *Electronic Journal of Statistics* **2**, 149-167.

12. **Zhang, H. H.** and Lu, W. (2007) Adaptive-LASSO for Cox's proportional hazard model. *Biometrika* **93**, 1-13.

13. Lu, W. and **Zhang, H. H.** (2007) Variable selection for linear transformation models via penalized marginal likelihood. *Statistics in Medicine* **26**, 3771-3781.

14. Leng, C. and **Zhang, H. H.** (2007) Nonparametric model selection in hazard regression. *Journal of Nonparametric Statistics* **18**, 417-429.

15. Liu, Y., **Zhang, H. H.**, Park, C. and Ahn, J. (2007) Support vector machines with adaptive L_q penalty. *Computational Statistics and Data Analysis*, **51**, 6380-6394.

16. **Zhang, H. H.** (2006) Variable selection for Support vector machines via smoothing spline ANOVA. *Statistica Sinica*, **16**, 659-674.

17. Lin, Y. and **Zhang, H. H.** (2006) Component selection and smoothing in smoothing spline analysis of variance models. *Annals of Statistics*, **34**, 2272-2297.

18. **Zhang, H. H.**, Ahn, J., Lin, X., and Park, C. (2006) Gene selection using support vector machines With Nonconvex Penalty. *Bioinformatics*, **22**, 88-95.

19. Liu, Y., **Zhang, H. H.**, Park, C. and Ahn, J. (2006) The L_q support vector machines. *Proceedings of Joint Summer Research Conference on Machine and Statistical Learning: Prediction and Discovery.* In press.

20. **Zhang, H. H.** and Lin, Y (2006) Component selection and smoothing for nonparametric regression in exponential families. *Statistica Sinica*, **16**, 1021-1042.

21. Tang, Y. and **Zhang, H. H.** (2005) Multiclass proximal support vector machines. *Journal of Computational and Graphical Statistics*, **15**, 339-355.

22. **Zhang, H. H.**, Wahba, G., Lin, Y., Voelker, M., Ferris, Klein, R. and Klein, B. (2004) Variable selection and model building via likelihood basis pursuit. *Journal of American Statistical Association* **99**, 659-672.

23. Ferris, M., Voelker, M. and **Zhang, H. H.** (2004) Model building with likelihood basis pursuit. *Journal of Optimization Methods and Software* **19**, 1-18.

24. Lin, Y., Wahba, G., **Zhang, H. H.**, and Lee, Y. (2002) Statistical properties and adaptive tuning of support vector machines. *Machine Learning* **48**, 115-136.

25. Wahba, G., Lin, Y., Lee, Y. and **Zhang, H. H.** (2003) Optimal properties and adaptive tuning of standard and nonstandard support vector machines. *Nonlinear Estimation and Classification*, Springer, 125-143.

26. Wahba, G., Lin, Y. and **Zhang, H. H.** (2000) Generalized approximate cross validation for support vector machines, or, another way to look at margin-like quantities. *Advances in Large Margin Classifiers*, MIT Press, 297-309.

Books

1. Clark, B., Fokoue, E., and **Zhang, H. H.** (2009) Principles and Theory for Data Mining and Machine Learning Series: Springer Series in Statistics.

Technical Reports/Submitted Papers

1. Zhang, H. H., Lu, W., and Wang, H. (2008) Penalized estimating equations for semiparametric linear transformation models. *submitted*.

2. Lu, W. and Zhang, H. H. (2008) On estimation of partially linear transformation models. *revised*.

3. Wahba, G., Lin, Y., Lee, Y. and Zhang, H. H. (2001) On the Relation between the GACV and Joachims' $\xi\alpha$ method for tuning Support Vector Machines, with extension to the nonstandard case. TR 1039, Department of Statistics, UW-Madison.

Research Support

1. National Science Foundation (NSF) CAREER Award
   Grant number: DMS-0645293, $400,000. 07/01/07 - 06/30/12
   Title: "Nonparametric Models Building, Estimation, and Selection with Applications to High Dimensional Data Mining."
   Role: PI

2. National Science Foundation (NSF) Standard Grant
   Grant number: DMS-0405913, $124,396. 07/01/04 - 06/30/07
   Title: "Nonparametric Variable Selection in Smoothing Spline ANOVA Models."
   Role: PI

3. National Institute of Health/National Cancer Institute (NIH/NCI)
   Grant number: R01 CA085848-08, $800,000. 05/01/07- 04/30/11
   Title: "Flexible Statistical Methods for Biomedical Data."
   Role: Co-investigator (20%)

4. NCSU Faculty Research and Professional Development Award Grant
   Grant number: 350609, $4,000. 03/01/03 - 02/29/04
   Title: "Nonparametric Variable Selection and Model Building."
   Role: PI

5. National Science Foundation (NSF)
   Grant number: DMS-0703392, $769,288. 07/01/07 - 06/30/12.
   Title: "CSUMS: NC State University Computation for Undergraduates in Statistics Program." **Role: Co-PI**

Travel Fund Support

1. NSF travel support, AMS/IMS/SIAM Joint Summer Research Conferences, 2006.
2. NIH travel support, ENAR conference, Junior Faculty Workshop, March 2004.
3. NSF travel support, Joint Statistical Meetings, August 2003.
4. NSF travel support, AMS/IMS/SIAM Joint Summer Research Conferences, 2003.
5. NSF travel support, IMS mini-meeting on Functional Data Analysis, January 2003.

Student Advising

- **Ph.D. Chair**

Lan Lan	graduated in May 2006 (co-advisor Daowen Zhang)
Xiao Ni	graduated in May 2007 (co-advisor Daowen Zhang)
Hongmei Yang	graduated in June 2007 (co-advisor Daowen Zhang)
Lingkang Huang	graduated in March 2008 (co-advisor Zhaobang Zeng)
Eren Demirhan	graduated in June 2008
Song Liu	graduated in July 2008
Justin Shows	graduated in 2009 (co-advisor Wenbin Lu)
Nan Li	expected to graduate in 2010
Mihye Ahn	expected to graduate in 2010 (co-advisor Wenbin Lu)
Dhruv Sharma	expected to graduate in 2010 (co-advisor Howard Bondell)

- **Master Chair**

Sangjin Park	graduated in May 2005
Elizabeth Nelson	graduated in May 2006
Karen Daughy	graduated in Sep 2006
John Schweitzer	graduated in March 2007

Professional Activities

- Program Committee, 13th International Conferences on Artificial Intelligence and Statistics (AISTAT), 2009.
- Regional Advisory Board, IBS/Eastern North American Region, 2004-2007
- Program Committee, 11th International Conferences on Artificial Intelligence and Statistics (AISTAT), 2007.
- Local Research Group on Data Mining, Group Leader, 2005-2006.
- Session Organizer and Chair, 2007 ICSA Applied Statistics Symposium
- Session Organizer and Chair, 2007 Nonparametric Statistics Conference
- Session Organizer and Chair, 2006 ENAR Spring Meetings.
- Session Chair, 2004 ENAR Spring Meeting,
- Session Chair, 2004 Joint Statistical Meetings, 2006 Joint Statistical Meetings
- Session Chair, 2003 Statistical Society of Canada Joint Meetings.

Invited Talks

1. "Nonparametric Variable Selection with COSSO and Its Recent Development", Colloquium, Department of Statistical Sciences, Cornell University, Ithaca, NY, Jan, 2009.
2. "Variable Selection for Multicategory SVM via Sup-Norm Regularization", Invited Talk at INFORMS Annual Meeting, Washington DC, May 2008.
3. "Model Selection for Partial Smoothing Splines", Invited Talk at Southern Regional Council on Statistics (SRCOS) Summer Research Conference, Charleston, SC, June 2008.
4. "Nonparametric Variable Selection with COSSO and Its Recent Development" , Invited Talk at International Conference on Advances in Interdisciplinary Statistics and Combinatorics , Greensboro, NC, October 2007.
5. "Nonparametric Variable Selection with COSSO and Its Recent Development" , Invited Talk at Southern Regional Council on Statistics (SRCOS), Richmond, VA, June 2007.
6. "Variable Selection in Linear Models via Shrinkage Methods", Colloquium, Department of Statistics, North Carolina State University, Raleigh, NC, Dec., 2006.

7. "Soft-thresholding Penalties for Variable Selection", Bayesian Group Seminar, Department of Statistics, NCSU, Raleigh, NC, Dec., 2006.

8. "Variable Selection via Penalized Likelihood Methods," Department Colloquium, Department of Statistics, University of Georgia, Athens, GA, Nov., 2006.

9. "Sparse Learning for Support Vector Machines", Invited Talk, Joint Statistical Meetings, Seattle, August, 2006.

10. "Support Vector Machine Classification with Informative Features", Invited Talk, AMS/IMS/SIAM Joint Summer Research Conferences – Machine Learning, Statistics and Discovery, June 2006, Snowbird, Utah.

11. "Support Vector Machine Classification with Informative Features", Invited Talk, INFORMS International Conference, June 2006, Hong Kong, China.

12. "Variable Selection for Support Vector Machines via COSSO," ISI Invited Talk, Joint Statistical Meetings, Minneapolis, MN, August, 2005.

13. "Variable Selection for SVM via Spline ANOVA," Invited Talk, The Joint Meeting of CSPS/IMS, Beijing, China, July 2005.

14. "Multiclass proximal support vector machines," Department Colloquium, Duke Clinical Research Institute (DCRI), Duke University, Durham, NC, May, 2005.

15. "Variable Selection via COSSO in Nonparametric Regression Models," Mathematisches Forschungsinstitut Oberwolfach workshop, Germany, Nov., 2004

16. "Component Selection and Smoothing Operator in Nonparametric Models," Department Colloquium, Department of Statistics, Pennsylvania State University, College Park, PA, Nov., 2004.

17. "Component Selection and Smoothing Operator in SS-ANOVA," Department Colloquium, Department of Biostatistics, University of North Carolina at Chapel Hill, Raleigh, NC, Sep., 2004.

18. "Variable Selection in Nonparametric Statistical Models," IMS Invited Talk, Joint Statistical Meetings, Toronto, Canada, August, 2004.

19. "Variable Selection in Nonparametric Statistical Models," Invited Talk, International Federation of Classification Societies meeting, Chicago, IL, July 2004.

20. "Unified Multi-class Proximal Support Vector Machines." Invited Talk, Interface: computational biology and bioinformatics, 36th Symposium on the interface, Baltimore, MA. May 2004.

21. "Component selection and smoothing in smoothing spline analysis of variance models." Contributed Talk, IBS-ENAR, Pittsburgh, PA. March 2004.

22. "Compactly Supported RBF Kernels." Statistical and Applied Mathematical Sciences Institute (SAMSI), Data Mining and Machine Learning Intermediate Workshop, Research Triangle Park, NC. Feb 2004.

23. "Compactly Supported Kernels", Poster, Statistical and Applied Mathematical Sciences Institute (SAMSI), Data Mining and Machine Learning workshop, Sep 8, 2003, Research Triangle Park, NC.

24. "Statistical Machine Learning", Department Colloquium, Department of Statistics, North Carolina State University, Raleigh, NC, August 19, 2003.

25. "Variable Selection in Spline ANOVA Models", Discussant for Wald Lecture (II) by Professor Grace Wahba, Joint Statistical Meetings, Aug 6, 2003, San Francisco, CA.

26. "Variable Selection for Support Vector Machines", IMS Invited Talk, Joint Statistical Meetings, Aug 4, 2003, San Francisco, CA.

27. "Nonparametric Regression Variable Selection – Component Selection and Smoothing Operator (COSSO)", Invited Talk, Statistical and Applied Mathematical Sciences Institute (SAMSI), Stochastic Computation Final Workshop, June 27, Research Triangle Park, NC.

28. "Variable Selection for Support Vector Machines via Basis Pursuit", Invited Talk, AMS/IMS/SIAM Joint Summer Research Conferences – Machine Learning, Statistics and Discovery, June 26, 2003, Snowbird, Utah.

29. "Nonparametric Variable Selection and Model Building via Likelihood Basis Pursuit", Poster, IMS mini-meeting on Functional Data Analysis, January 9-10, 2003, Gainesville, FL.

30. "Nonparametric Variable Selection via Likelihood Basis Pursuit", Colloquium, Department of Statistics, Duke University, Durham, September 6, 2002.

31. "Nonparametric Variable Selection via Likelihood Basis Pursuit with Applications in Medical Studies", Contributed Talk, ENAR, Nonparametric Section, Arlington, VA, March 2002.

32. "Variable Selection via Basis Pursuit for non-Gaussian data", Department Colloquium, Department of Statistics, University of Wisconsin-Madison, Madison, November 7, 2001.

33. "Variable Selection via Basis Pursuit for non-Gaussian data", Contributed Talk, Joint Statistical Meetings, Atlanta, August 2001.

References

Available upon request