

Lan Huong Nguyen

PhD Candidate · ICME, Stanford University

27 Olmsted Rd Apt 105, Stanford, CA 94305 · (+1) 626-200-7792 · lanhuong@stanford.edu

🏠 <http://nlhuong.github.io> | 💻 github.com/nlhuong | 🔗 www.linkedin.com/in/lan-huong-nguyen

Education

Stanford University

Stanford, CA

PhD in Computational and Mathematical Engineering (GPA 3.9)

Sep. 2013 - Exp. Jun. 2019

- Robust non-linear dimensionality reduction and latent structure recovery. Advised by Prof. Susan Holmes.
- Relevant coursework: numerical linear algebra, convex optimization, stochastic processes, probabilistic graphical models, distributed algorithms, statistical learning theory, Bayesian statistics, algorithms in advanced machine learning, theories of deep learning, statistical and machine learning methods for genomics

California Institute of Technology

Pasadena, CA

B.S. in Applied Mathematics and Economics (GPA 3.8)

Sep. 2009 - Jun. 2013

- Graduated with honors from the Department of Computing & Mathematical Sciences.

Experience

High-dimensional Data Statistics Graduate Research

Stanford, CA

Researcher in Holmes Lab

Sep. 2014 - Present

- Design robust dimensionality reduction and manifold learning methods for inferring latent structures and associated uncertainties in high-dimensional data
- Build statistical models for single cell RNA-seq and microbiome (16S, metagenomics, and metatranscriptomics) data
- Develop **R** and **Shiny** software packages implementing new data analysis and visualization techniques

R for Data Science Introductory Course

Stanford, CA

Course Instructor for CME/STATS 195

Fall '16, '17, '18

- Created a short course on R application to data science covering topics such as data wrangling, exploration, visualization and modeling
- Taught a class of 60+ graduate and undergraduate students
- Course details and syllabus can be found at: <http://cme195.github.io>

Genentech

South San Francisco, CA

Intern in Department of Bioinformatics

Jun. 2017 - Oct. 2017

- Studied the role of COP1 in activation and repression kinetics of LPS-induced genes
- Developed an open-source Bioconductor package for analyzing and visualizing short time-course RNA-seq data

AOL Advertising

Palo Alto, CA

Intern in Large-Scale Analytics Team

Jun. 2014 - Oct. 2014

- Analyzed DoubleClick ad exchange auction bid data consisting of ~4-5 million entries per day
- Identified factors affecting winning rates of bids and developed classification models for predicting successful bidding
- Estimated market trends and winning price distribution. Designed Real Time Bidding algorithms

Talentoday Startup

Palo Alto, CA

Consultanting Project

Sep. 2012 - Feb. 2013

- Worked with 30k+ user data to find the most predictive survey questions using QR factorization, and sparse CCA.
- Developed an answer-imputation technique using binomial matrix completion.
- Reduced the number of questions by 30% while retaining high test accuracy.

Skills

Programming R, Stan, Python, Matlab, C++, Spark, Scala, SQL

Languages English, Polish, Vietnamese

Honors & Awards

ICME Stanford University

Senior Teaching Fellow

Institute of Computational and Mathematical Engineering teaching award.

Stanford, CA

2017-2018

University of Chicago Midwest Trading Competition

First Place

Chicago, IL

Apr. 2013

The UChicago Midwest Trading Competition is the nation's premier algorithmic trading competition. Participants are asked to develop algorithms to make automated trading decisions. The competition tests technical programming skills, financial acumen, and ability to adjust based on algorithm performance.

Publications and Packages

L.H. Nguyen, S. Holmes. **Diffusion t-SNE for manifold learning**. Neurips (under review)

M. Roodgar, A. Babveyh, L. H. Nguyen, W. Zhou, R. Sinha, H. Lee, J. B. Hanks, M. Avula, L. Jiang, H. Lee, G. Song, H. Chaib, I. L. Weissman, S. Batzoglou, S. Holmes, D. G. Smith, J. L. Mankowski, S. Prost, M. P. Snyder. **Chromosome-level de novo assembly of the pig-tailed macaque genome using linked-read sequencing and HiC proximity scaffolding**. Nature Methods (under review)

J.A. Grembi, L.H. Nguyen, T.D. Haggerty, C.D. Gardner, S.P. Holmes and J. Parsonnet. **Gut microbiota plasticity correlated with sustained weight loss after a low-carb or low-fat dietary intervention**. Scientific Reports (under review)

L.H. Nguyen, S. Holmes. **Ten quick tips for effective dimensionality reduction**. PLOS Comp. Biology (2019) <https://doi.org/10.1371/journal.pcbi.1006907>

L.H. Nguyen., S. Holmes. **Bayesian Unidimensional Scaling for visualizing uncertainty in high dimensional datasets with latent ordering of observations**. BMC Bioinformatics (2017) <https://doi.org/10.1186/s12859-017-1790-x>

L.H. Nguyen. **'TimeSeriesExperiment' package for analysis and visualization of short time course data**. Bioconductor (2018) <https://doi.org/doi:10.18129/B9.bioc.TimeSeriesExperiment>

P.H.T. Kamga, B. Li, M. McKerns, L.H. Nguyen, M. Ortiz, H. Owhadi, and T.J. Sullivan. **Optimal uncertainty quantification with model uncertainty and legacy data**. Journal of the Mechanics and Physics of Solids, 72 (2014) <http://dx.doi.org/10.1016/j.jmps.2014.07.007>

Presentations

BioVis at International Society for Computational Biology

Prague, Czech Republic, Jul. 2017

Presented an accepted paper on Bayesian dimensionality reduction for recovering laternt ordering of samples.

Ascona Workshop 2017: Statistical Challenges in Single-Cell Biology

Ascona, Switzerland, Apr. 2017

Presented a poster on mathematics of t-SNE and diagnostics for choosing its hyperparameters.

ASA's Women in Statistics and Data Science

Charlotte, NC, Oct. 2016

Presented methods for detecting gradients in microbiome data based on indirect ordination. Invited student panelist.

useR 2016

Stanford, CA, Jun. 2016

Presented **mvarVis**: An R package for Visualization of Multivariate Analysis Results.

Bioinformatics for Microbiome Symposium

Stanford, CA, May 2016

Presented a comparative study of differential abundance methods for analyzing sequencing data.

Additional Teaching Experience

Stanford, CA, Oct. 2014 - Present

Teaching assistant for: BIOS 221/STATS 366 Modern Statistics for Modern Biology, CS 261 Optimization and Algorithmic Paradigms, CME 200 Linear Algebra with Application to Engineering Computations, CME 106 Intro to Probability and Statistics

Extracurricular Activities _____

Stanford Women in Mathematics, Statistics and Computational Engineering Co-Founder and Financial officer	Stanford, CA Mar. 2018 - Present
Stanford Polish Student Association Co-president, financial officer	Stanford, CA Oct. 2016 - Jun. 2018
Asian American Graduate Students Association Core Member	Stanford, CA Jun. 2015 - Jun. 2016