Sam Roweis, 1972–2010

Computer scientist and statistical astronomer Sam Roweis took his own life in New York City on 2010 January 12. He was a brilliant and accomplished researcher in the field of machine learning, and a strong advocate for the use of computational statistics for automating discovery and scientific data analysis. He made several important contributions to astronomy and was working on adaptive astronomical data analysis at the time of his death.

Roweis obtained his PhD in 1999 from the California Institute of Technology, where he worked on a remarkable range of subjects, including DNA computing, modeling of dynamical systems, signal processing, and speech recognition. During this time he unified and clarified some of the most important data analysis techniques, including Principal Component Analysis, Hidden Markov Models, and Expectation Maximization. His work was aimed at making data analysis and modeling faster, but also better justified scientifically. The last years of his PhD were spent in Princeton NJ, where he came in contact with a young generation of cosmologists thinking about microwave background and large-scale structure data.

In a postdoc at University College London, Roweis co-created the Locally Linear Embedding (LLE) algorithm; a simple but flexible technique for mapping a large data set onto a low-dimensional manifold. The LLE paper obtained more than 2700 citations in 9 years, launched a new sub-field of machine learning known as "manifold learning", and inspired work in data visualization, search, and applied mathematics.

In 2001, Roweis took a faculty job at the University of Toronto Computer Science Department. He continued working on data analysis methods that have probabilistic interpretation and therefore scientific applicability, but at the same time have good performance on large data sets. He was awarded a Sloan Fellowship, a Canada Research Chair, and a fellowship of the Canadian Institute for Advanced Research, among other honors and awards. During this period he turned some of his attention to astronomy, beginning with a project to infer a distribution function (in this case the velocity distribution of disk stars) in the face of non-trivial measurement errors and missing data. He also contributed inference ideas and optimization technology to a precise photometric re-calibration of the *Sloan Digital Sky Survey*. He was awarded tenure at Toronto in 2006.

In 2005, Roweis began the *Astrometry.net* collaboration. Under his leadership, this group built software that can take any astronomical image of

unknown provenance and rapidly and robustly determine its pointing, orientation, and plate scale with no first guess or prior information of any kind. The system works in part by converting astrometric calibration into a well-posed problem in decision theory. It is now recovering corrupted data in plate-scanning projects, calibrating data taken by amateur astronomers, and working inside the photo-sharing site *flickr*. *Astrometry.net* was the primary PhD project of Roweis's student Dustin Lang (now at Princeton University), who is one of the first PhDs in computer science to obtain a postdoc in astronomy. These interdisciplinary successes demonstrated the enormous potential of having Roweis thinking about astronomical data.

In September 2009, after a few years at search giant *Google*, he took a tenured position in the New York University Computer Science Department. At the time of his death, he was working on flexible data analysis systems for imaging and spectroscopy, capitalizing on and contributing to NYU's involvement in *SDSS-III*; the idea was that pipelines should not just *reduce* the science data, they should also simultaneously *learn* instrument and calibration parameters from the union of the calibration and science data.

In his astronomical projects he advocated straightforward and simple models that are flexible and therefore live in large parameter spaces; these are feasible only with good engineering. The *Astrometry.net* system works by brute force search, after geometric hashing has trimmed the tree of possibilities by about 15 orders of magnitude. His data-reduction pipelines usually had more parameters than data; these are incredibly flexible for automated discovery of instrument properties but they require clever regularization.

Many students came to machine learning after inspiration from Roweis, and many astronomers modified their techniques and approaches after even short conversations with him. His extremely popular on-line video lectures give beautiful examples of his clear and engaging style. He was an ideal collaborator: reliable, funny, enthusiastic, creative, and outrageously intelligent.

He is survived by his father Shoukry Roweis, his wife Meredith Goldwasser, and his two daughters Aya and Orli.

David W. Hogg
New York University