

SON Series

Amundson Lecture Series 2009 Abstracts Presented by: Professor Émmanuel Candes Wednesday, April 29th General Colloquium (UH Hilton: Flamingo Room #275) : Compressive Sensing Thursday, April 30th Seminar Lecture (TLC2: Philip G. Hoffman Hall #232): Recovering the Unseen: Is the Netflix Problem Well Posed? Friday, May 1st Graduate Student Lecture (Philip G. Hoffman Hall #347) : L1-magic

Compressive Sensing

Recovering the Unseen: * Is the Netflix Problem Well Posed?

One of the central tenets of signal processing and data acquisition is the Shannon/Nyquist sampling theory: the number of samples needed to capture a signal is dictated by its bandwidth. This talk introduces a novel sampling or sensing theory which goes against this conventional wisdom. This theory now known as "Compressed Sensing or Compressive Sampling" allows the faithful recovery of signals and images from what appear to be highly incomplete sets of data, i.e. from far fewer measurements or data bits than traditional methods use. We will present the key ideas underlying this new sampling or sensing theory, and will survey some of the most important results. We will emphasize the practicality and the broad applicability of this technique, and discuss what we believe are far reaching implications; e.g. procedures for sensing and compressing data simultaneously and much faster. Finally, there are already many ongoing efforts to build a new generation of sensing devices based on compressed sensing and we will discuss remarkable recent progress in this area as well.

This talk considers a problem of considerable practical interest: the recovery of a data matrix from a sampling of its entries. In partially filled out surveys, for instance, we would like to infer the many missing entries. In the area of recommender systems, users submit ratings on a subset of entries in a database, and the vendor provides recommendations based on the user's preferences. Because users only rate a few items, we would like to infer their preference for unrated items (this is the famous Netflix problem). Formally, suppose that we observe m entries selected uniformly at random from a matrix. Can we complete the matrix and recover the entries that we have not seen?

We show that perhaps surprisingly, one can recover low-rank matrices exactly from what appear to be highly incomplete sets of sampled entries; that is, from a comparably small number of entries. Further, perfect recovery is possible by solving a simple convex optimization program, namely, a convenient semidefinite program (SDP). This result hinges on powerful techniques in probability theory. We will also present a very efficient algorithm based on iterative singular value thresholding, which can complete matrices with about a billion entries in a matter of minutes on a personal computer. In many applications, one often has fewer equations than unknowns. While this seems hopeless, the premise that the object we wish to recover is sparse or nearly sparse radically changes the problem, making the search for solutions feasible. This lecture will introduce sparsity as a key modeling tool, and introduce a series of little miracles touching on many areas of data processing. These examples show that finding that solution to an underdetermined system of linear equations with minimum L1 norm, often returns the "right" answer.