Maximum Likelihood Degree and PI

Source: http://sci.tech−archive.net/Archive/sci.stat.math/2004−07/0199.html

---

From: Osher Doctorow (*mdoctorow_at_comcast.net*)
Date: 07/15/04

```
Date: Thu, 15 Jul 2004 00:21:04 +0000 (UTC)
```

From Osher Doctorow mdoctorow@comcast.net

F. Catanese, S. Hosten, A. Khetan, and B. Sturmfels, in "The maximum likelihood degree," arXiv:math.AG/0406533 v1 25 Jun 2004, discuss some of the latest developments in Algebraic Statistics, which intersects quite a bit with my recent threads here.

The "AG" in the arXiv listing refers to Algebraic Geometry, which readers should look up in the Front for the Mathematics ArXiv as keywords under Algebraic Geometry, and goes back to the early 1990s where some of the pioneering papers appeared. It is also interesting as an exercise to see who came up with original ideas and who made "variations on a theme," although some people who do variations on a theme eventually learn to develop original ideas later on under certain circumstances.

Three papers cited by Catanese et al are L. Garcia, M. Stillman, and B. Sturmfelds, "Algebraic geometry of Bayesian networks," math.AG/ 0301255, J. Symbolic Comput., to appear; L. Pachter and B. Sturmfels, "Tropical geometry of statistical models, q−bio.QM/0311009, Proc. Natl. Acad. Sci. USA, to appear; and G. Pistone, E. Riccomagno, and H. P. Qynn, "Algebraic statistics: computational commutative algebra in statistics," Chapman and Hall, Boca Raton, Florida, 2001. How− ever, some of the older references cited have involved very Creative people who have remained at the top of the field, and readers should look at the paper themselves. Of course, where there are Bayesian networks, a PI interpretation cannot be far behind (see above).

Readers will find themselves introduced to or looking up concepts like sheaves, degrees of algebraic functions, generating functions, independence models, irreducible projective varieites, logarithmic singularities, normal crossing divisors, Chern class, projective d−space, smooth toric varieties, Newton polytopes, bounded regions, arrangements of hyperplanes, linear forms, resolution of singularit− ies, etc. The hyperplanes {f_i = 0} are reminiscent of PI express− ions, and the authors derive a formula for the Maximum Likelihood (ML) degree when the f_i's are Laurent polynomials, as for example:

1) f_i = ai + bi(theta_1) + gi(theta_2) + di(theta_1)(theta_2)

for i = 1, 2, 3, 4, where i is a subscript and _ also represents a subscript and the polynomials are in variables theta_1 and theta_2 with coefficients bi, gi, di. The authors hope that algorithms can eventually be developed to better solve the critical equations dlog(f) = 0 for f = f_1^u1 f_2^u2 ... f_n^un and so on.

The authors point out that their paper was motivated by recent papers on ML degree in statistics and computational biology, including a paper on phylogenetic models in 2003, a paper on unidirected graphical models in 2002, a paper on ML degree of a Gaussian graphical model (2004), a paper on ML degree of mixture models (2004), where specific references to these are listed in their reference section. For all these papers, the exact numerical values of the ML degree were found.

So what is the ML degree? It is the degree of the algebraic function that is an optimal solution theta−hat to a problem which is an algebraic function of the data u, namely the problem of maximizing f_1(theta)^u1 f_(theta)^u2...f_n(theta)^un for theta in an appropriate subset of R^d. The paper begins with discrete data and later does some work on certain continuous data.

Osher Doctorow