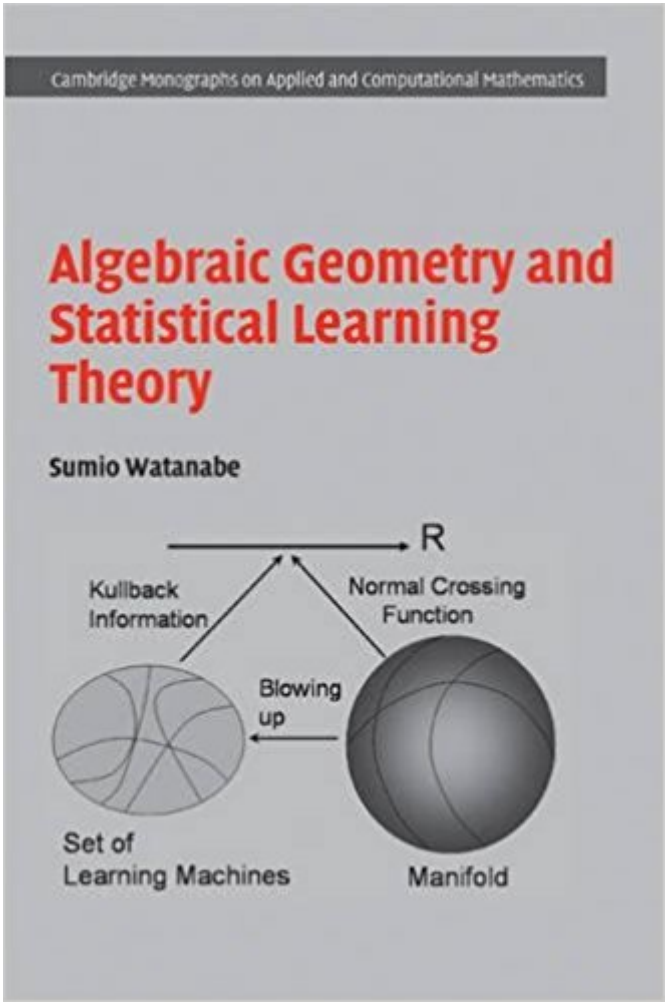
# Algebraic Geometry And Statistical Learning Theory (Cambridge Monographs On Applied And Computational Mathematics)



## Synopsis

Sure to be influential, Watanabe’s book lays the foundations for the use of algebraic geometry in statistical learning theory. Many models/machines are singular: mixture models, neural networks, HMMs, Bayesian networks, stochastic context-free grammars are major examples. The theory achieved here underpins accurate estimation techniques in the presence of singularities.

## Book Information

Series: Cambridge Monographs on Applied and Computational Mathematics (Book 25)
Hardcover: 300 pages
Publisher: Cambridge University Press; 1 edition (September 28, 2009)
Language: English
ISBN-10: 0521864674
ISBN-13: 978-0521864671
Product Dimensions: 6 x 0.8 x 9 inches
Shipping Weight: 1.2 pounds (View shipping rates and policies)
Average Customer Review: 4.4 out of 5 stars 2 customer reviews
Best Sellers Rank: #630,027 in Books (See Top 100 in Books) #101 inÃ Â Books > Science & Math > Mathematics > Geometry & Topology > Algebraic Geometry #127 inÃ Â Books > Computers & Technology > Computer Science > AI & Machine Learning > Computer Vision & Pattern Recognition #372 inÃ Â Books > Textbooks > Science & Mathematics > Mathematics > Geometry

## Customer Reviews

"Overall, the many insightful remarks and simple direct language make the book a pleasure to read." >

Sure to be influential, this book lays the foundations for the use of algebraic geometry in statistical learning theory. Many widely used statistical models are singular: mixture models, neural networks, HMMs, and Bayesian networks are major examples. The theory achieved here underpins accurate estimation techniques in the presence of singularities.

Statistical learning theory is now a well-established subject, and has found practical use in artificial intelligence as well as a framework for studying computational learning theory. There are many fine books on the subject, but this one studies it from the standpoint of algebraic geometry, a field which

decades ago was deemed too esoteric for use in the real world but is now embedded in myriads of applications. More specifically, the author uses the resolution of singularities theorem from real algebraic geometry to study statistical learning theory when the parameter space is highly singular. The clarity of the book is outstanding and it should be of great interest to anyone who wants to study not only statistical learning theory but is also interested in yet another application of algebraic geometry. Readers will need preparation in real and functional analysis, and some good background in algebraic geometry, but not necessarily at the level of modern approaches to the subject. In fact, the author does not use algebraic geometry over algebraically closed fields (only over the field of real numbers), and so readers do not need to approach this book with the heavy machinery that is characteristic of most contemporary texts and monographs on algebraic geometry. The author devotes some space in the book for a review of the needed algebraic geometry.Also reviewed in the initial sections of the book are the concepts from statistical learning theory, including the very important method of comparing two probability density functions: the Kullback-Leibler distance (called relative entropy in the physics literature). The reader will have to have a good understanding of functional analysis to follow the discussion, being able to appreciate for example the difference between convergence in different norms on function space. From a theoretical standpoint, learning can be different in different norms, a fact that becomes readily apparent throughout the book (from a practical standpoint however, it is difficult to distinguish between norms, due to the finiteness of all data sets). Of particular importance in early discussion is the need for "singular" statistical learning theory, which as the author shows, boils down to finding a mathematical formalism that can cope with learning problems where the Fisher information matrix is not positive definite (in this case there is no guarantee that unbiased estimators will be available). This is where (real) algebraic geometry comes in, for it allows the removal of the singularities in parameter space by recursively using "blow-up" (birational) maps. The author lists several examples of singular theories, such as hidden Markov models, Boltzmann machines, and Bayesian networks. The author also shows to generalize some of the standard constructions in "ordinary" or "regular" statistical learning to the case of singular theories, such as the Akaike information criterion and Bayes information criterion. Some of the definitions he makes are somewhat different than what some readers are used to, such as the notion of stochastic complexity. In this book it is defined merely as the negative logarithm of the `evidence', whereas in information theory it is a measure of the code length of a sequence of data relative to a family of models. The methods for calculating the stochastic complexity in both cases are similar of course.In singular theories, one must deal with such things as the divergence of the maximum likelihood estimator and the failure of asymptotic

normality. The author shows how to deal with these situations after the singularities are resolved, and he gives a convincing argument as to why his strategies are generic enough to cover situations where the set of singular parameters, i.e. the set where the Fisher information matrix is degenerate, has measure zero. In this case, he correctly points out that one still needs to know if the true parameter is contained in the singular set, and this entails dealing with "non-generic" situations using hypothesis testing, etc.Examples of singular learning machines are given towards the end of the book, one of these being a hidden Markov model, while another deals with a multilayer perceptron. The latter example is very important since the slowness in learning in multilayer perceptrons is widely encountered in practice (largely dependent on the training samples). The author shows how this is related to the singularities in the parameter space from which the learning is sampled, even when the true distribution is outside of the parametric model, where the collection of parameters is finite. This example leads credence to the motto that "singularities affect learning" and the author goes on further to show to what extent this is a "universal" phenomenon. By this he means that having only a "small" number of training samples will bring out the complexity of the singular parameter space; increasing the number of training samples brings out the simplicity of the singular parameter space. He concludes from this that the singularities make the learning curve smaller than any nonsingular learning machine. Most interestingly, he speculates that "brain-like systems utilize the effect of singularities in the real world."

Today program development is guided by best practices, traditions, and downright doctrinal beliefs.Machine learning in particular has always been a set of magical techniques, especially when applied to linguistic problems.Pr Watanabe is paving the way to a world where it is an actual science.This is one of the most important books I know for the future of computer science engineering.The problem is that reading this requires ÃƒÂ  bernerd level mathematics background *and* mindset (do not wait for too much pedagogy). So chances are, if you are of mathematician breed you will love it ; if your cursus is computer science engineering, it willl make you cry.

Geometry and Field Theory (Cambridge Monographs on Mathematical Physics) Commutative Algebra: with a View Toward Algebraic Geometry (Graduate Texts in Mathematics) Algebraic Geometry (Graduate Texts in Mathematics) Algebraic Geometry: A First Course (Graduate Texts in Mathematics) (v. 133) Applied Computational Aerodynamics: A Modern Engineering Approach (Cambridge Aerospace Series) Differential Equations and Their Applications: An Introduction to Applied Mathematics (Texts in Applied Mathematics) (v. 11) Principles of Mathematical Analysis (International Series in Pure and Applied Mathematics) (International Series in Pure & Applied Mathematics) Introduction to the Foundations of Applied Mathematics (Texts in Applied Mathematics) Nelson Pure Mathematics 2 and 3 for Cambridge International A Level (Nelson Mathematics for Cambridge International a Level) Algebraic Graph Theory (Graduate Texts in Mathematics) Algebraic Topology: An Intuitive Approach (Translations of Mathematical Monographs, Vol. 183) An Introduction to Scientific Computing: Twelve Computational Projects Solved with MATLAB (Texts in Applied Mathematics) Computational Inelasticity (Interdisciplinary Applied Mathematics) (v. 7) Probability on Trees and Networks (Cambridge Series in Statistical and Probabilistic Mathematics) Random Graphs and Complex Networks: Volume 1 (Cambridge Series in Statistical and Probabilistic Mathematics) Data Analysis and Graphics Using R: An Example-Based Approach (Cambridge Series in Statistical and Probabilistic Mathematics)