Why

We use a loss function and a predictive density to construct a regressor on the domain of the random process.

Setup

Let $\ell : \mathbf{R} \times \mathbf{R} \to \mathbf{R}$ be a loss function. We choose a predictor to minimize the expected loss under the predictive density.

Squared error case

Consider $\ell(\alpha, \beta) = (\alpha - \beta)^2$. The *minimum squared error normal random function predictor* or *minimum squared error gaussian process predictor* for dataset $(a^1, \gamma_1), \ldots, (a^n, \gamma_n)$ in $A \times \mathbf{R}$ is the predictor which minimizes the squared error loss. Since the predictive density is normal, the minimizer is the conditional mean.

Absolute error case

Consider $\ell(\alpha, \beta) = |\alpha - \beta|$. The *minimum absolute deviation normal random function predictor* or *minimum absolute deviation gaussian process predictor* for dataset $(a^1, \gamma_1), \ldots, (a^n, \gamma_n)$ in $A \times \mathbf{R}$ is the predictor which minimizes the absolute deviation loss. For any density, the solution is the median. Since the predictive density is normal, and so symmetric, the median is the conditional mean. In other words, the minimum absolute deviation normal random function predictor coinicides with the minimium squared error normal random function predictor.

Definition

For this reason, the *normal random function predictor* or *gaussian process predictor* for dataset $(a^1, \gamma_1), \ldots, (a^n, \gamma_n)$ in $A \times \mathbf{R}$ is $h : A \to \mathbf{R}$ defined by

$$h(x) = m(x) + \Big(k(x, a^1) \cdots k(x, a^n)\Big) \left(\Sigma_a + \Sigma_e\right)^{-1} (\gamma - m_a).$$

In other words, the regressor which assigns to each point its conditional mean.

Notice that h is an affine function of γ. If the mean function $m \equiv 0$ then h is linear in γ. This is sometimes called a *linear estimator*.[1] Alternatively, notice (in the zero mean setting) that h is a linear combination of n kernel function $k(x, a^i)$ for $i = 1, \ldots, n$. Specifically, h is a linear combination of

The process of using a normal random function predictor is often called *Gaussian process regression* or (especially in spatial statistics) *kriging*. The upside is that a gaussian process predicor interpolates the data, is smooth, and the so-called variance increases with the distance from the data.[2]

[1] We avoid the other terminology we have seen used—linear predictor—because the predictor h is not linear in its input x.

[2] This last piece is true for certain covariance kernels and will be clarified in future editions.


Normal Random Function Regressors
Normal Random Function Predictive Densities
Affine Transformations
Normal Random Functions
Multivariate Normals
Random Functions
Transformations
Positive Definite Matrices
Symmetric Real Matrices
Quadratic Forms
Normal Densities
Multivariate Real Densities
Matrix Trace
Real Matrix Inverses
Probability Densities
Matrix Transpose
Real Matrix-Matrix Products
Exponential Function
Real Integrals
Real Matrix-Vector Products
Random Vectors
Nonnegative Integrals
Matrix Determinants
Vectors as Matrices
Real Linear Combinations
Simple Integrals
Real Inner Product
Real Matrices
Random Variables
Vectors
Real Norm
Linear Equations
Multivariate Real Polynomials
Measurable Functions
Probability Measures
Topological Sigma Algebra
Loss Functions
Space Inner Product
Linear Functions
Arrays
Real Vectors
Multivariate Functions
Measures
Topologies
Similarity Functions
Predictor Performance Metrics
Plane Inner Product
Space Norm
N-Dimensional Space
Space Vectors
Plane Vector Sums and Differences
Sigma Algebras
Metrics
Real Square Roots
Right Triangle Sides Relation
Plane Norm
Distance
Real Series
Extended Real Numbers
Real Squares
Plane Vectors
Space Distance
Real Polynomials
Real Limits
Event Probabilities
Simple Functions
Plane Distance
Absolute Value
Real Arithmetic
Probability Distributions
Real Functions
Interval Length
Real Multiplicative Inverses
Outcome Probabilities
Predictors
Natural Square Roots
Intervals
Real Products
Real Space
Real Summation
Real Sequences
Square Numbers
Real Line
Real Additive Inverses
Real Plane
Natural Summation
Sequences
Subset Algebras
Squares
Permutations
Integral Line
Real Order
Real Sums
Rational Arithmetic
Rational Multiplicative Inverses
Lists
Cardinality
Fields
Groups
Geometry
Rational Order
Real Numbers
Rational Sums
Rational Products
Set Numbers
Inverse Elements
Rings
Rational Numbers
Finite Sets
Integer Arithmetic
Element Functions
Direct Products
Integer Products
Natural Multiplicative Identity
Natural Additive Identity
Equivalent Sets
Natural Equations
Comparisons
Integer Order
Integer Sums
Arithmetic
Identity Elements
Function Inverses
Family Operations
Uncertain Events
Family Unions and Intersections
Orders
Equations
Integer Numbers
Natural Powers
Function Images
Function Composites
Operations
Families
Natural Order
Converse Relations
Natural Products
Equivalence Relations
Functions
Natural Sums
Relations
Recursion Theorem
Ordered Pair Projections
Uncertain Outcomes
Peano Axioms
Natural Induction
Characteristic Functions
Set Products
Subset Systems
Natural Numbers
Set Powers
Generalized Set Dualities
Ordered Pairs
Partitions
Successor Sets
Intersection of Empty Set
Unordered Triples
Set Symmetric Differences
Set Dualities
Set Unions and Intersections
Set Intersections
Pair Unions
Set Complements
Set Differences
Set Unions
Pair Intersections
Empty Set
Unordered Pairs
Set Specification
Set Equality
Set Inclusion
Deductions
Quantified Statements
Logical Statements
Statements
Sets
Identities
Names
Letters
Objects