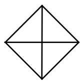


Why

We use a normal linear model to predict the function at inputs not included in the design.

Definition

Let $(x : \Omega \to \mathsf{R}^d, A \in \mathsf{R}^{n\times d}, e : \Omega \to \mathsf{R}^n)$ be a normal linear model over the probability space $(\Omega, \mathcal{A}, \mathbf{P})$. A *predictive linear model* is a linear model with two additional objects. The first is a matrix $B \in \mathsf{R}^{m\times d}$ and the second is a random vector $f : \Omega \to \mathsf{R}^m$ which is jointly normal with e. So a predictive linear model is a sequence (x, A, e, C, f).

As usual we define a random vector $y : \Omega \to \mathsf{R}^n$, but now also define a random vector $z : \Omega \to \mathsf{R}^m$ by

$$
\begin{aligned}
y &= Ax + e \\
z &= Cx + f.
\end{aligned}
$$

The *predictive density* of the predictive normal linear model is the conditional density of z given y. Observe that

$$
\begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} I & 0 & 0 \\ A & I & 0 \\ C & 0 & I \end{pmatrix} \begin{pmatrix} x \\ e \\ f \end{pmatrix}.
$$

and the vector (x, e, f) is normal with covariance

$$
\begin{pmatrix} \Sigma_x & 0 & 0 \\ 0 & \Sigma_e & \Sigma_{ef} \\ 0 & \Sigma_{fe} & \Sigma_f \end{pmatrix}
$$

So (x, y, z) is normal with covariance

$$\begin{pmatrix} \Sigma_x & \Sigma_x A^\top & \Sigma_x C^\top \\ A\Sigma_x & A\Sigma_x A^\top + \Sigma_e & A\Sigma_x C^\top + \Sigma_{ef} \\ C\Sigma_x & C\Sigma_x A^\top + \Sigma_{fe} & C\Sigma_x C^\top + \Sigma_f \end{pmatrix}$$

A *normal linear model predictor* or *normal linear model regressor* is the predictor which assigns to a new point $a \in \mathbf{R}^d$ the mean of the predictive density. In other words, the predictor $g : \mathbf{R}^d \to \mathbf{R}$ defined by

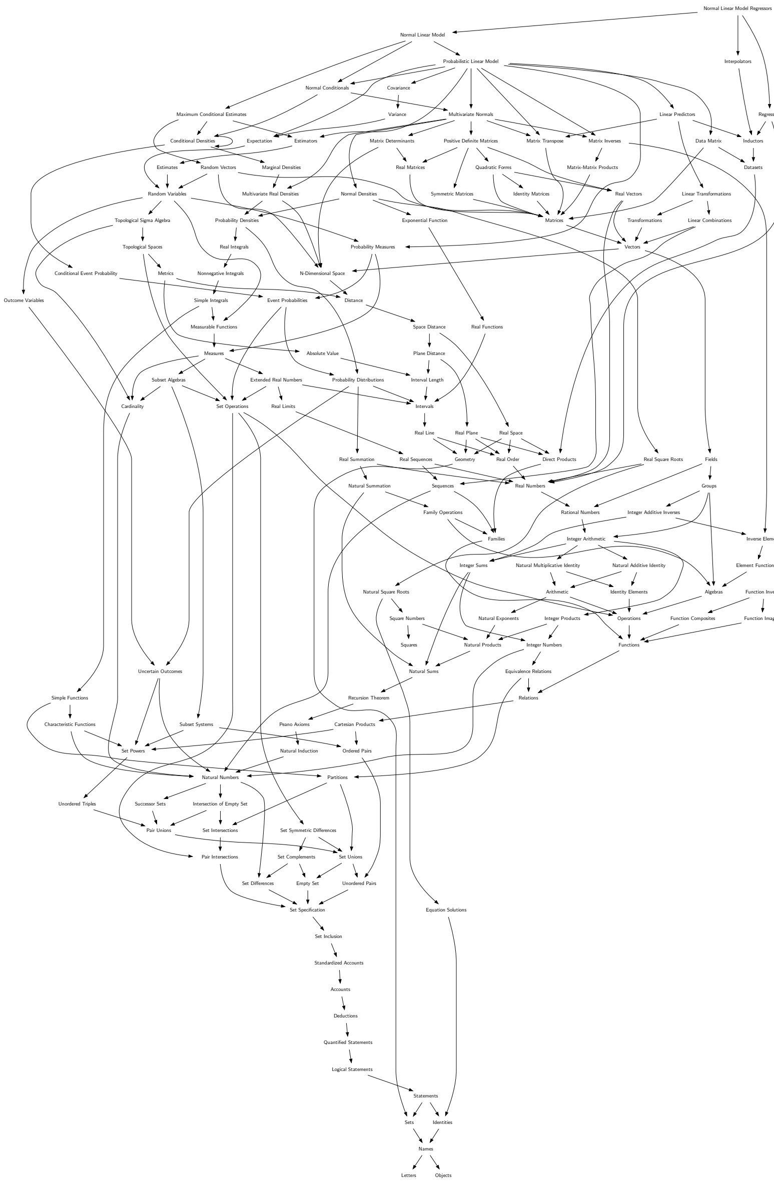
$$g(a) = (a^\top \Sigma_x A^\top + \Sigma_{fe}) \left(A\Sigma_x A^\top + \Sigma_e\right)^{-1} \gamma.$$

In the above we have substituted $a^\top$ for C. In the case of normal random vectors this corresponds with the MAP esetimate and the MMSE estimate.[1] The predictive density for this point is normal with that mean, and has variance

$$(a^\top \Sigma_x a + \Sigma_f) - (a^\top \Sigma_x A^\top + \Sigma_{fe}) \left(A\Sigma_x A^\top + \Sigma_e\right)^{-1} (A\Sigma_x a + \Sigma_{ef}).$$

Use of a predictive normal linear model is often referred to as *Gaussian process regression*. 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. This is also called *Bayesian linear regression*.

[1] Future editions will have discussed this and include a reference to the sheet.


Normal Linear Model Regressors
Normal Linear Model
Probabilistic Linear Model
Interpolators
Normal Conditionals
Covariance
Maximum Conditional Estimates
Variance
Multivariate Normals
Linear Predictors
Regressors
Conditional Densities
Expectation
Estimators
Matrix Determinants
Positive Definite Matrices
Matrix Transpose
Matrix Inverses
Data Matrix
Inductors
Estimates
Random Vectors
Marginal Densities
Real Matrices
Quadratic Forms
Matrix-Matrix Products
Datasets
Random Variables
Multivariate Real Densities
Normal Densities
Symmetric Matrices
Identity Matrices
Real Vectors
Linear Transformations
Topological Sigma Algebra
Probability Densities
Exponential Function
Matrices
Transformations
Linear Combinations
Topological Spaces
Real Integrals
Probability Measures
Vectors
Conditional Event Probability
Metrics
Nonnegative Integrals
N-Dimensional Space
Outcome Variables
Simple Integrals
Event Probabilities
Distance
Measurable Functions
Space Distance
Real Functions
Measures
Absolute Value
Plane Distance
Subset Algebras
Extended Real Numbers
Probability Distributions
Interval Length
Cardinality
Set Operations
Real Limits
Intervals
Real Line
Real Plane
Real Space
Real Summation
Real Sequences
Geometry
Real Order
Direct Products
Real Square Roots
Fields
Natural Summation
Sequences
Real Numbers
Groups
Family Operations
Rational Numbers
Integer Additive Inverses
Families
Integer Arithmetic
Inverse Elements
Integer Sums
Natural Multiplicative Identity
Natural Additive Identity
Element Functions
Natural Square Roots
Arithmetic
Identity Elements
Algebras
Function Inverses
Square Numbers
Natural Exponents
Integer Products
Operations
Function Composites
Function Images
Squares
Natural Products
Integer Numbers
Functions
Uncertain Outcomes
Natural Sums
Equivalence Relations
Simple Functions
Recursion Theorem
Relations
Characteristic Functions
Subset Systems
Peano Axioms
Cartesian Products
Set Powers
Natural Induction
Ordered Pairs
Natural Numbers
Partitions
Unordered Triples
Successor Sets
Intersection of Empty Set
Pair Unions
Set Intersections
Set Symmetric Differences
Pair Intersections
Set Complements
Set Unions
Set Differences
Empty Set
Unordered Pairs
Set Specification
Equation Solutions
Set Inclusion
Standardized Accounts
Accounts
Deductions
Quantified Statements
Logical Statements
Statements
Sets
Identities
Names
Letters
Objects