Why

We want an estimator for the parameters of a linear function, given observations of the function with additive noise.

Definition

Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space. Let $x : \Omega \to \mathbf{R}^d$. Define $g : \Omega \to (\mathbf{R}^d \to \mathbf{R})$ by $g(\omega)(a) = a^\top x(\omega)$, for $a \in \mathbf{R}^d$. In other words, for each outcome $\omega \in \Omega$, $g_\omega : \mathbf{R}^d \to \mathbf{R}$ is a linear function with parameters $x(\omega)$. g_ω is the function of interest.

Let $a^1, \ldots, a^n \in \mathbf{R}^d$ a dataset with data matrix $A \in \mathbf{R}^{n \times d}$. Let $e : \Omega \to \mathbf{R}^n$ independent of x, and define $y : \Omega \to \mathbf{R}^n$ by

$$y = Ax + e.$$

In other words, $y_i = x^\top a^i + e_i$.

We call (x, A, e) a *probabilistic linear model*. Other terms include *linear model*, *linear regression model*, *bayesian linear regression*, and *bayesian analysis of the linear model*.[1] We call x the parameters, A a *design*, e the *error* or *noise* vector, and y the *observation* vector.

One may want an estimator for the parameters x in terms of y or one may be modeling the function g and want to predict $g(a)$ for $a \in A$ not in the dataset.

[1] The word bayesian is in reference to treating the object of interest—x—as a random variable.

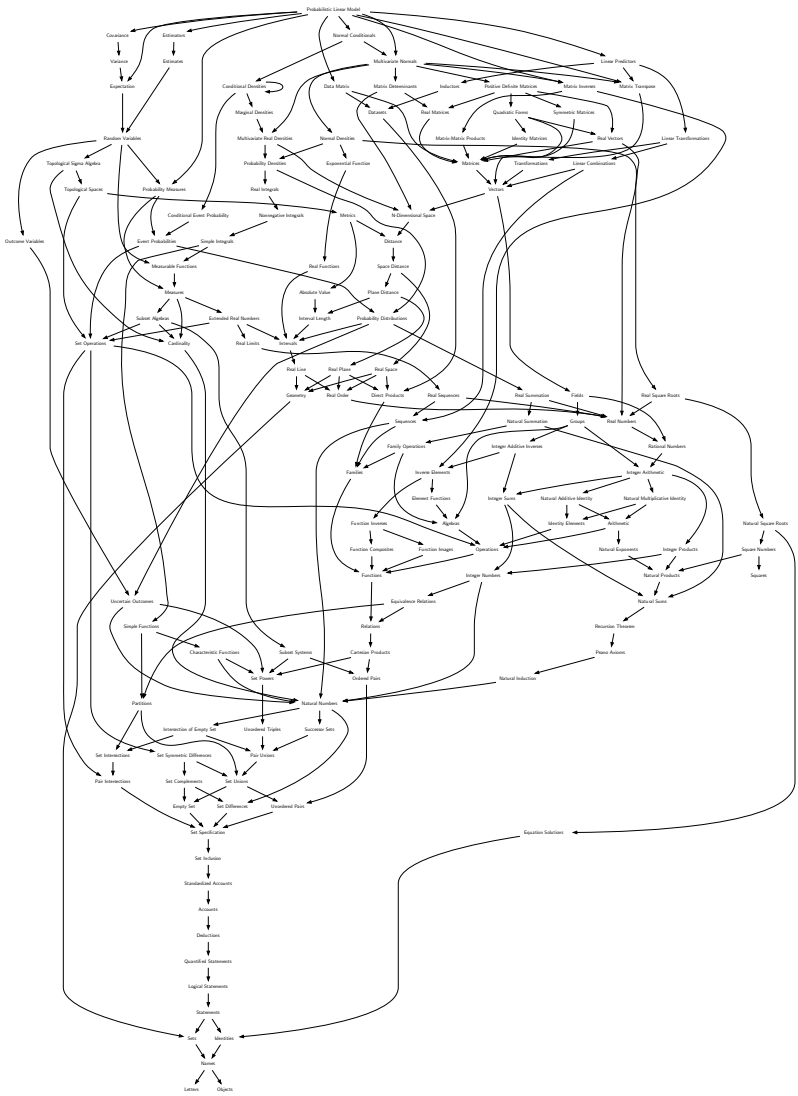
Mean and variance

Proposition 1. $\mathsf{E}(y) = A\,\mathsf{E}(x) + \mathsf{E}(w)$[2]

Proposition 2. $\mathsf{cov}((x, y)) = A\,\mathsf{cov}(x)A^\top + \mathsf{cov}\,e$[3]

[2]By linearity. Full account in future editions.

[3]Full account in future editions.


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