Why

What is the best linear predictor if we choose according to a squared loss function.

Definition

Let $X \in \mathbf{R}^{n \times d}$ and $y \in \mathbf{R}^d$. In other words, we have a paired dataset of records with inputs in $\mathbf{R}^d$ (the rows of X) and outputs in $\mathbf{R}$ (the elements of y).

A *least squares linear predictor* or *linear least squares predictor* is a linear transformation $f : \mathbf{R}^d \to \mathbf{R}$ (the field is $\mathbf{R}$) which minimizes

$$\frac{1}{n} \sum_{i=1}^{n} (f(x^i) - y_i)^2.$$

over the dataset of pairs $(x^1, y_1), \ldots, (x^n, y_n) \in \mathbf{R}^d \times \mathbf{R}$where x^i is the ith row of X for $i = 1, \ldots, n$.

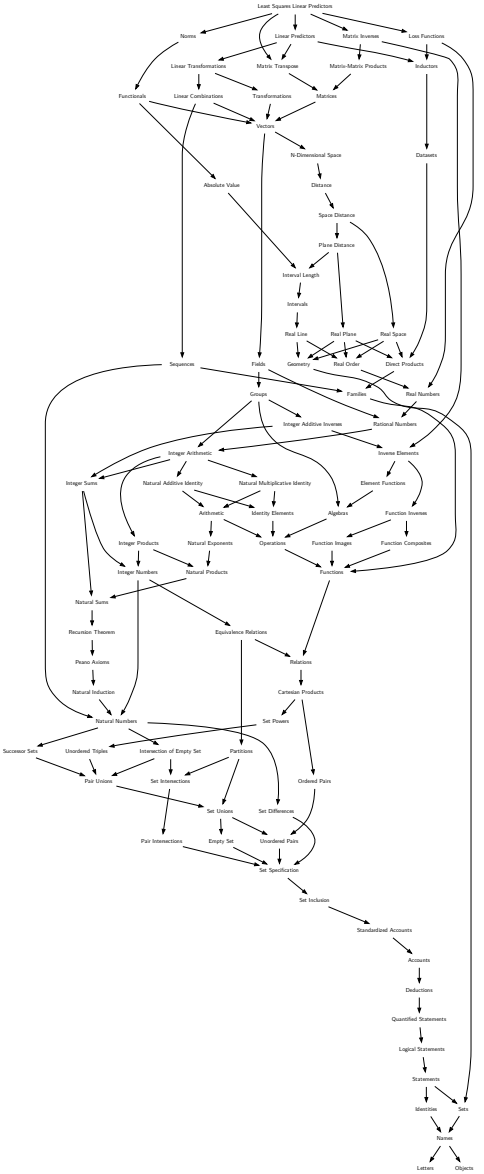
The set of linear functions from $\mathbf{R}^d$ to $\mathbf{R}$ is in one-to-one correspondence with $\mathbf{R}^d$. So we want to find $\theta \in \mathbf{R}^d$ to minimize

$$\frac{1}{n} \|X\theta - y\|^2.$$

Solution

Proposition 1. *There exists a unique linear least squares predictor and its parameters are given by* $\left(X^\top X\right)^{-1} X^\top y$.[1]

[1] Future editions will include an account.


Least Squares Linear Predictor
Norms
Linear Predictors
Matrix Inverses
Loss Functions
Linear Transformations
Matrix Transpose
Matrix-Matrix Products
Inductors
Functionals
Linear Combinations
Transformations
Matrices
Vectors
N-Dimensional Space
Datasets
Absolute Value
Distance
Space Distance
Plane Distance
Interval Length
Intervals
Real Line
Real Plane
Real Space
Sequences
Fields
Geometry
Real Order
Direct Products
Families
Real Numbers
Groups
Integer Additive Inverses
Rational Numbers
Integer Arithmetic
Inverse Elements
Integer Sums
Natural Additive Identity
Natural Multiplicative Identity
Element Functions
Arithmetic
Identity Elements
Algebras
Function Inverses
Integer Products
Natural Exponents
Operations
Function Images
Function Composites
Integer Numbers
Natural Products
Functions
Natural Sums
Recursion Theorem
Equivalence Relations
Peano Axioms
Relations
Natural Induction
Cartesian Products
Natural Numbers
Set Powers
Successor Sets
Unordered Triples
Intersection of Empty Set
Partitions
Pair Unions
Set Intersections
Ordered Pairs
Set Unions
Set Differences
Pair Intersections
Empty Set
Unordered Pairs
Set Specification
Set Inclusion
Standardized Accounts
Accounts
Deductions
Quantified Statements
Logical Statements
Statements
Identities
Sets
Names
Letters
Objects