Why

What is the best linear predictor if we choose according to a particular norm.

Definition

Suppose we have a paired dataset of n records with inputs in $\mathbf{R}^d$ and outputs in $\mathbf{R}$. A *norm weighted least squares linear predictor* for a norm $g : \mathbf{R}^n \to \mathbf{R}$ is a linear transformation $f : \mathbf{R}^d \to \mathbf{R}$ (the field is $\mathbf{R}$) which minimizes

$$g(y - Ax).$$

Weight matrix

Let $\|\cdot\|_W$ be the weighted norm for some positive semidefinite weight matrix W. We want to find x to minimize

$$\|y - AX\|_W.$$

This problem is referred to by many authors as *weighted least squares* or the *weighted least squares problem*.

Diagonal weight matrix

A special case of norm weighted least squares with a weighted norm is the usual weighted least squares problem (see Weighted Least Squares Linear Predictors). Consider weighted least squares with weights $w \in \mathbf{R}^n$, $w \geq 0$. Define $W \in \mathbf{R}^{n \times n}$ so that $W_{ii} = w_i$ and $W_{ij} = 0$ when $i \neq j$. So, in particular, W is a diagonal matrix and

$$\|y - Ax\|_W = \sum_{i=1}^{n} w_i (y_i - x^\top a_i)^2.$$

Solution

Proposition 1. *There exists a unique weighted least squares linear predictor and its parameters are given by*

$$(A^\top W A)^{-1} A^\top W y.$$

Norm Weighted Least Squares Linear Regressors

Weighted Least Squares Linear Regressors

Least Squares Linear Regressors

Linear Prediction

Real Function Approximation — Matrix Transpose — Loss Function — Input Designs

Real Matrix Inverses — Predictor Performance Metrics — Consistent Induction — Regressors

Real Matrix-Matrix Products — Predictors

Vector Space Bases

Span

Linearly Dependent Vectors — Subspaces — Real Matrix-Vector Products

Norms — Linear Transformation — Real Subspaces — Real Inner Products — Vector as Matrices — Real Matrices

Functionals — Real Function Space — Transformations — Linear Combination — Real Vector Span — Real Inner Product — Real Norm — Linear Equations

Similarity Functions — Vectors — Real Linear Combinations — Space Inner Product — Linear Functions — Approximation

Metrics — N-Dimensional Space — Plane Vector Sums and Differences — Space Vector — Plane Norm — Real Vectors — Space Norm — Arrays

Absolute Value — Right Triangle Sides Relation — Distance — Plane Inner Product — Plane Vectors

Real Square — Real Functions — Space Distance

Real Arithmetic — Plane Distance

Interval Length — Real Multiplication Inverse

Intervals — Real Space — Real Products

Square Numbers — Real Line — Integral Line — Real Plane — Real Additive Inverse — Real Sums — Real Summation — Real Optimizers — Lists

Squares — Rational Arithmetic — Real Order — Rational Multiplicative Inverse — Real Numbers

Geometry — Rational Products — Rational Order — Fields — Rational Sums — Inverse Elements — Set Number — Direct Products

Rings — Rational Numbers — Optimizers — Element Functions — Finite Sets

Integer Arithmetic — Natural Additive Identity — Natural Multiplicative Identity — Natural Summation — Equivalent Sets — Function Inverses — Family Unions and Intersections

Integer Products — Natural Equations — Integer Order — Arithmetic — Identity Elements — Comparison — Family Operations — Function Images — Function Composite — Families

Integer Numbers — Natural Powers — Equations — Orders — Operations — Functions

Natural Products — Natural Order — Converse Relations — Equivalence Relations

Natural Sums — Relations

Recursion Theorem

Peano Axioms — Ordered Pair Projections

Natural Induction — Set Products

Natural Numbers — Set Powers — Generalized Set Dualities

Set Symmetric Differences — Successor Sets — Unordered Triples — Set Dualities — Intersection of Empty Set — Partitions — Set Unions and Intersections

Set Complements — Pair Unions — Ordered Pairs — Set Intersections

Set Differences — Set Unions — Unordered Pairs — Pair Intersections

Empty Set — Set Specification

Set Equality

Set Inclusion

Deductions

Quantified Statements

Logical Statements

Statements

Sets — Identities

Names

Letters — Objects