Why

We model a real-valued output as corrupted by small random errors. Thus, we can talk about a dataset which is "close" to being consistent with a linear predictor.

Definition

Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space. Let $x \in \mathbf{R}^d$ and $e : \Omega \to \mathbf{R}^n$. For $A \in \mathbf{R}^{n \times d}$, define $y : \Omega \to \mathbf{R}^n$ by $y = Ax + e$. We call (x, A, e) a *probabilistic errors linear model*. We call y the *response vector*, A the *model matrix* and e the *error vector*.

Moment assumptions

The most basic distributional assumption for a probabilistic errors linear model pertain to the expectation and variance. Since $\mathbf{E}(y) = Ax + \mathbf{E}(e)$ and $\mathbf{var}(y) = \mathbf{var}(e)$, these assumptions can be given for e or for y.

If $\mathbf{E}(x) = 0$ and $\mathbf{var}(y) = \sigma^2 I$ then we call (x, A, e) a *classical linear model with moment assumptions*. Notice that the components of e are assumed uncorrelated. We have $d + 1$ unknowns (the $d \times 1$ entires of θ and scalar parameter σ^2.

In this case $\mathbf{E}(y_i) = {a^i}^\top \theta$ and so θ is called the *mean parameter vector* and σ^2 is called the *model variance*. The model variance indicates the variability inherent in the observations. Neither the mean nor variance of the error depends on the regression vector x nor on the parameter vector θ.

Examples

Consider the *two-sample problem* in which we have two populations with (unknown) mean responses $\alpha_1, \alpha_2 \in \mathbf{R}$. We observe these responses with (perhaps unknown) common variance σ^2, and assume that errors are uncorrelated.

We define $y^1 = \alpha_1 \mathbf{1} + e^1$ and $y^2 = \alpha_2 \mathbf{1} + e^2$ so that we can stack these and obtain

$$y = \begin{bmatrix} y^1 \\ y^2 \end{bmatrix} = \begin{bmatrix} \alpha_1 \mathbf{1} \\ \alpha_2 \mathbf{1} \end{bmatrix} + \begin{bmatrix} e^1 \\ e^2 \end{bmatrix}.$$

To cast this in our standard form we define

$$A = \begin{bmatrix} \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \cdots & \begin{bmatrix} 1 \\ 0 \end{bmatrix} & \begin{bmatrix} 0 \\ 1 \end{bmatrix} & \cdots & \begin{bmatrix} 0 \\ 1 \end{bmatrix} \end{bmatrix}^\top, \quad x = \begin{bmatrix} \alpha_1 \\ \alpha_2 \end{bmatrix}.$$

with regression vectors $a_1 = (1, 0)$ and $x_2 = (0, 1)$ repeated n_1 and n_2 times, respectively. An input design for this model involves specifying a sequence of these two vectors, which (with the uncorrelated assumption) reduces to dictating how many responses should be collected from each population. The inputs here is really the set $\mathcal{X} = \{1, 2\}$. The feature function is $\phi : \mathcal{X} \to \mathbf{R}^2$ defined by $\phi(1) = (1, 0)$ and $\phi(2) = (0, 1)$. And so the regression range is $\{(1, 0), (0, 1)\}$.


Probabilistic Ames Linear Model
Feature Maps
Covariance
Random Variable Moments
Least Squares Linear Regression
Linear Predictors
Real Matrix Inverses
Loss Functions
Real Function Maximization
Input Design
Real Matrix Matrix Products
Vector Space Bases
Prediction Performance Metrics
Consistent Induction
Algorithms
Probabilistic Predictors
Matrix Transpose
Span
Linearly Dependent Vectors
Predictors
Zero Matrix
Subspaces
Real Matrix Vector Products
Vectors
Vectors as Matrices
Matrices
Real Subspaces
Linear Transformations
Expectation
Functionals
Real Vectors
Real Vector Spaces
Transformations
Linear Combinations
Random Variable
Real Inner Product
Real Linear Combinations
Topological Sigma Algebra
Measurable Functions
Probability Measures
Real Norm
Linear Regression
Matrices
Vectors
Topological Spaces
Outcome Variable Covariance
Measures
Linear Functions
Matrix Functions
N-Dimensional Space
Metrics
Space Vectors
Plane Vector Sums and Differences
Outcome Variable Expectation
Sigma Algebra
Plane Inner Product
Space Point
Plane Through Origin Predicate
Vectors
Real Square Roots
Induced Probability Distributions
Real Series
Extended Real Numbers
Plane Norm
Space Distance
Real Squares
Real Limits
Outcome Variable Events
Absolute Value
Plane Distance
Plane Vectors
Real Addends
Sums Probabilities
Interval Length
Arrays
Real Multiplication Inverses
Real Functions
Probability Distributions
Real Means
Real Powers
Integers
Natural Square Roots
Real Sequences
Real Negatives
Squares
Real Order
Rational Multiplication Inverses
Natural Addends
Real Sums
Real Quotients
Integral Lists
Real Summations
Integers
Natural Order
Rational Products
Rational Sums
Real Numbers
Rings
Income Maximize
Subset Algebras
Does Products
Set Numbers
Natural Numbers
Cardinality
Commutes
Integer Multiplication
Finite Sets
Inverse Functions
Probable Outcomes
Comparisons
Integer Order
Integer Products
Natural Addition Identity
Natural Multiplication Identity
Equivalent Sets
Natural Summation
Family Unions and Intersections
Divides
Integer Pairs
Natural Pairs
Natural Orders
Equalities
Natural Powers
Function Images
Injective Completions
Identities
Families
Integer Quotients
Concrete Relations
Equivalence Relations
Functions
Natural Sums
Recursive Theorems
Set Products
Subset Systems
Set Powers
Ordered Pairs
Name Scopes
Natural Induction
Unordered Triples
Uncurried Set Duplets
Natural Numbers
Set Unions and Intersections
Set Duplets
Intersection of Empty Set
Successor Sets
Partitions
Set Symmetric Differences
Pair Unions
Set Complements
Pair Intersections
Set Differences
Set Unions
Empty Set
Unordered Pairs
Set Equalities
Set Equality
Set Inclusion
Standardized Axiomatic
Axiomatic
Deductions
Quantified Statements
Logical Statements
Statements
Sets
Identities
Names
Leaves
Objects