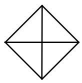


Why

We consider the probabilistic linear model in which all random variables involved are normal.

Definition

The *normal linear model* is a linear model in which the signal and noise have normal (Gaussian) densities. For this reason, the model is often called the *Gaussian linear model* or the *linear model with Gaussian noise.*

Let $(\Omega, \mathcal{A}, \mathbf{P})$ be a probability space. Let $x : \Omega \to \mathsf{R}^d$ and $e : \Omega \to \mathsf{R}^d$ be independent normal random vectors with zero mean and covariances Σ_x and Σ_e. Let $y = Ax + e$. We have n precepts in R^d. So let $a^1, \dots, a^n \in \mathsf{R}^d$ with data matrix $A \in \mathsf{R}^{n \times d}$.

Maximum conditional estimate of x

Since x and e are gaussian, y is gaussian. So the random vector (x, y) has normal density with mean zero and covariance

$$\begin{pmatrix} \Sigma_x & \Sigma_x A^\top \\ A\Sigma_x & A\Sigma_x A^\top + \Sigma_e \end{pmatrix}.$$

This is also called *bayesian linear regression* or the *bayesian analysis of the linear model.* The word bayesian is in reference to treating the quantity of interest—x—as a random variable.

Proposition 1. *The maximum conditional estimate of* x :

$\Omega \to \mathsf{R}^d$ given observed value $\gamma \in \mathsf{R}^n$ of $y : \Omega \to \mathsf{R}^n$ is the conditional mean $\Sigma_{xy}{\Sigma_{yy}}^{-1}\gamma$.

Recall that the maximum conditional estimate also maximizes the joint density.

Uncorrelated noise

Suppose that $\Sigma_e = \sigma^2 I$.

Proposition 2. *A solution to maximize $g(\alpha, \gamma)$ with respect to α is $\alpha = -\Sigma^{-1}\Sigma X^\top \gamma$.*

Proposition 3. *$g_{\theta|y}(\alpha, \gamma)$ is normal with mean*

$$\tilde{\mu}(\gamma) = \Sigma X^\top \left(X \Sigma X^\top\right)^{-1} \gamma$$

and covariance

$$\tilde{\Sigma} = \Sigma - \Sigma X^\top \left(X \Sigma X^\top\right)^{-1} X \Sigma.$$

Proposition 4. *A solution to maximize $g_{\theta|y}(\alpha, \gamma)$ w.r.t. α is*

$$\tilde{\Sigma}\tilde{\Sigma}^{-1}\tilde{\mu}(\gamma).$$

But, of course, y also has a density. Denote the density of y by $g : \mathsf{R}^n \to \mathsf{R}$. In other words, $g \geq 0$ and $\int g = 1$.

Proposition 5.

$$\log g(\gamma) = -1/2(\gamma^\top \left(X \Sigma X^\top\right)^{-1} \gamma) - \frac{d}{2} \log 2\pi - \frac{1}{2} \log \mathbf{det} \left(X \Sigma X^\top\right)$$

Test

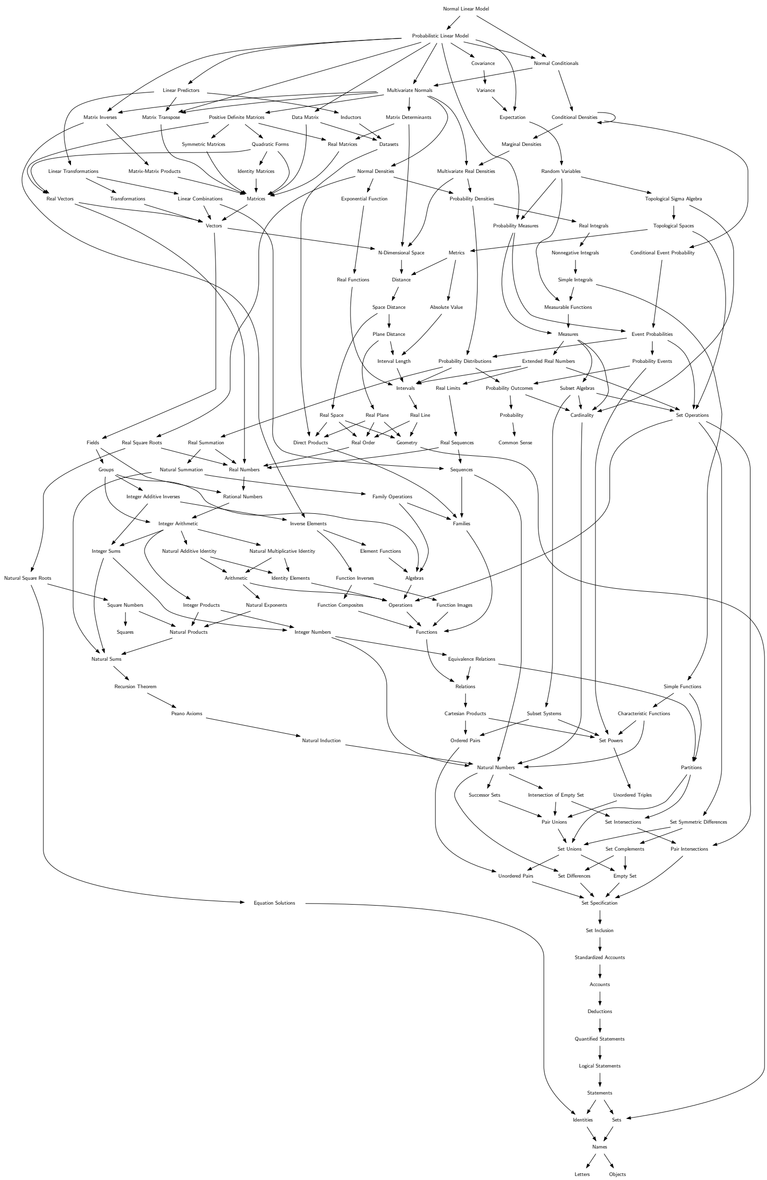
This expression makes clear that y is has a normal density with mean $X\,\mathsf{E}(x)$ and covariance $X\,\mathsf{E}(x)X^\top$.

Let $w : \Omega \to \mathsf{R}^d$ be a random vector with mean 0 and covariance ηI. Let $x^1, \dots, x^n \in \mathsf{R}^d$ Define $y^i : \Omega \to \mathsf{R}$ by $y_i(\omega) = w(\omega)^\top x^i$ for $i = 1, \dots, d$.

Noise setup

Let $e : \Omega \to \mathsf{R}^n$ be a normal random vector with mean 0 and covariance σI. Define $\tilde{y} : \Omega \to \mathsf{R}^n$ by $\tilde{y} = y(\omega) + e(\omega)$.

Proposition 6. *$\tilde{y}$ is a normal random vector with mean zero and covariance $X \Sigma X^\top + \sigma I$.*


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