Why

[1]

Definition

We want to estimate a random variable $x : \Omega \to \mathbf{R}^n$ from a random variable $y : \Omega \to \mathbf{R}^n$ using an estimator $\phi : \mathbf{R}^m \to \mathbf{R}^n$ which is affine.[2] In other words, $\phi(\xi) = A\xi + b$ for some $A \in \mathbf{R}^{n\times m}$ and $b \in \mathbf{R}^n$. We will use the mean squared error cost.

We want to find A and b to minimize

$$\mathbf{E}\|Ax + b - y\|^2.$$

Proof. Express $\mathbf{E}(\|Ax + b - y\|^2)$ as $\mathbf{E}((Ax + b - y)^\top(Ax + b - y))$

$$\begin{array}{ccccccc} + & \operatorname{tr}(A\mathbf{E}(xx^\top)A^\top) & + & \mathbf{E}(x)^\top A^\top b & - & \operatorname{tr}(A^\top\mathbf{E}(yx^\top)) \\ + & b^\top A\mathbf{E}(x) & + & b^\top b & - & b^\top\mathbf{E}(y) \\ - & \operatorname{tr}(A\mathbf{E}(xy^\top)) & - & \mathbf{E}(y)^\top b & + & \mathbf{E}(yy^\top) \end{array}$$

The gradients with respect to b are

$$\begin{array}{ccccccc} + & 0 & + & A\mathbf{E}(x) & - & 0 \\ + & A\mathbf{E}(x) & + & 2b & - & \mathbf{E}(y) \\ - & 0 & - & \mathbf{E}(y) & + & 0 \end{array}$$

so $2A\mathbf{E}(x) + 2b - 2\mathbf{E}(y)$. The gradients with respect to A are

$$\begin{array}{ccccccc} + & \mathbf{E}(xx^\top)A^\top + \mathbf{E}(xx^\top)^\top A^\top & + & \mathbf{E}(x)b^\top & - & \mathbf{E}(yx^\top)^\top \\ + & \mathbf{E}(x)b^\top & + & 0 & - & 0 \\ - & \mathbf{E}(xy^\top) & - & 0 & + & 0 \end{array}$$

so $2\mathbf{E}(xx^\top)A^\top + 2\mathbf{E}(x)b^\top - 2\mathbf{E}(xy^\top)$. We want A and b solutions to

$$\begin{aligned} A\mathbf{E}(x) + b - \mathbf{E}(y) &= 0 \\ \mathbf{E}(xx^\top)A^\top + \mathbf{E}(x)b^\top - \mathbf{E}(xy^\top) &= 0 \end{aligned}$$

[1] Future editions will include an account.

[2] Actually, the development flips this. Future editions will correct.

so first get $b = \mathbf{E}(y) - A\mathbf{E}(x)$. Then express

$$\begin{aligned}
\mathbf{E}(xx^\top)A^\top + \mathbf{E}(x)(\mathbf{E}(y) - A\mathbf{E}(x))^\top - \mathbf{E}(xy^\top) &= 0. \\
\mathbf{E}(xx^\top)A^\top + \mathbf{E}(x)\mathbf{E}(y)^\top - \mathbf{E}(x)\mathbf{E}(x)^\top A^\top - \mathbf{E}(xy^\top) &= 0. \\
(\mathbf{E}(xx^\top) - \mathbf{E}(x)\mathbf{E}(x)^\top)A^\top &= \mathbf{E}(xy^\top) - \mathbf{E}(x)\mathbf{E}(y)^\top. \\
\mathbf{cov}(x,x)A^\top &= \mathbf{cov}(x,y).
\end{aligned}$$

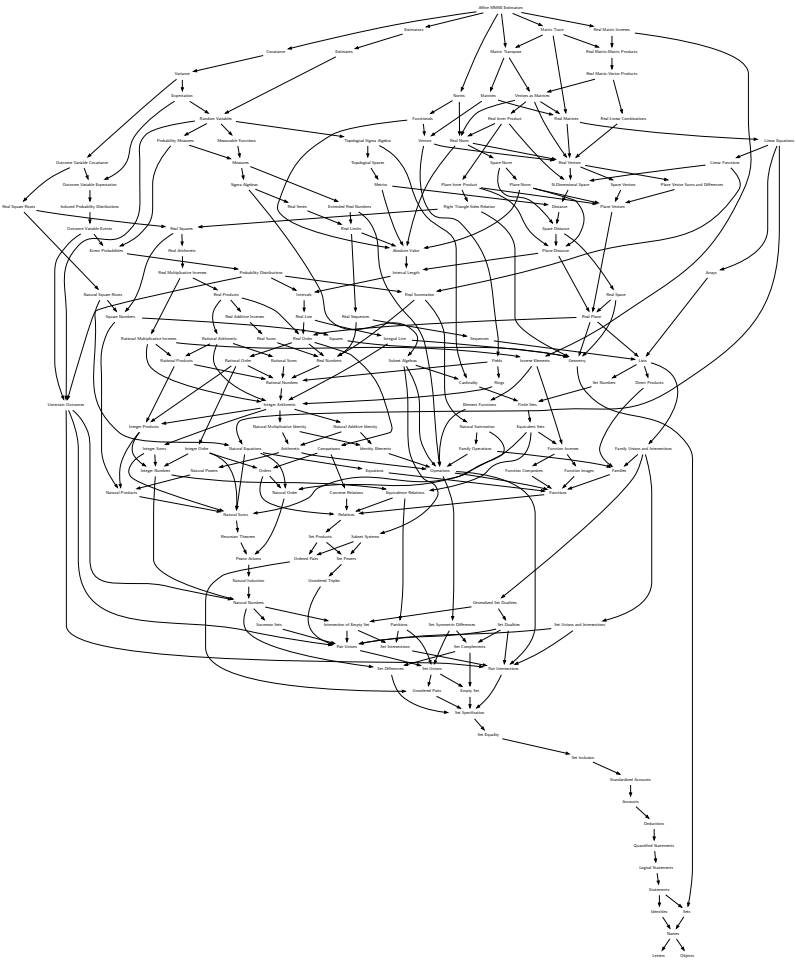
So $A^\top = \mathbf{cov}(x,x)^{-1}\mathbf{cov}(x,y)$ means $A = \mathbf{cov}(y,x)\mathbf{cov}(x,x)^{-1}$ is a solution. Then $b = \mathbf{E}(y) - \mathbf{cov}(y,x)\mathbf{cov}(x,x)^{-1}\mathbf{E}(x)$. So to summarize, the estimator $\phi(x) = Ax + b$ is

$$\mathbf{cov}(y,x)\left(\mathbf{cov}x,x\right)^{-1}x + \mathbf{E}(y) - \mathbf{cov}(y,x)\mathbf{cov}(x,x)^{-1}\mathbf{E}(x)$$

or

$$\mathbf{E}(y) + \mathbf{cov}(y,x)\left(\mathbf{cov}x,x\right)^{-1}(x - \mathbf{E}(x))$$

$\square$


Mean MSME Estimators
Estimators
Covariance
Estimates
Matrix Transpose
Matrix Trace
Real Matrix Inverses
Real Matrix-Matrix Products
Real Matrix-Vector Products
Vectors
Expectation
Norms
Matrices
Vectors as Matrices
Random Variable
Functionals
Real Norm Product
Real Matrices
Real Linear Combinations
Probability Measures
Measurable Functions
Topological Space Algebra
Metrics
Real Spaces
Linear Equations
Outcome Variable Covariance
Measures
Topological Spaces
Space Norms
Real Metrics
Linear Functions
Outcome Variable Expectation
Sigma Algebra
Metrics
Plane Inner Product
Plane Norm
Metrical Space
Space Metrics
Plane Vector Norms and Differences
Real Square Roots
Indicator Probability Distributions
Real Series
Extended Real Numbers
Right Triangle Sides Relation
Distance
Plane Metrics
Outcome Variable Events
Real Squares
Real Limits
Space Distance
Event Probabilities
Real Arithmetic
Absolute Value
Plane Distance
Real Multiplication Inverses
Probability Distributions
Interval Length
Sums
Natural Square Roots
Real Products
Integrals
Real Summations
Real Sums
Square Numbers
Real Addition Inverses
Real Line
Real Sequences
Real Power
Rational Multiplication Inverses
Rational Multiplications
Real Sums
Real Lines
Sigma
Integral Lines
Sequences
Sets
Rational Products
Rational Order
Rational Sums
Real Numbers
Subset Algebras
Fields
Interior Elements
Sums
Rational Numbers
Cardinality
Sets
Set Numbers
Disjoint Products
Unordered Statements
Integer Arithmetic
Element Functions
Finite Sets
Integer Products
Natural Multiplicative Identity
Natural Additive Identity
Natural Summations
Equivalent Sets
Integer Sums
Integer Order
Natural Equations
Addition
Commutative
Bijective Mappings
Family Operations
Finite Sums
Family Unions and Intersections
Integer Numbers
Natural Powers
Power
Equivalence
Mappings
Function Compositions
Function Images
Cardinal
Natural Products
Natural Sums
Concrete Relations
Equivalence Relations
Functions
Natural Sums
Relations
Recursive Theorems
Set Products
Subset Systems
Finite Sums
Ordered Pairs
Set Power
Natural Induction
Unordered Triples
Natural Numbers
Generalized Set Modifiers
Successor Sets
Intersection of Empty Set
Functions
Set Symmetric Differences
Set Modifiers
Set Unions and Intersections
Pair Unions
Set Intersections
Set Complements
Set Differences
Set Unions
Pair Intersections
Unordered Pairs
Empty Set
Set Specifications
Set Equality
Set Inclusion
Standardized Accounts
Accounts
Deductions
Quantified Statements
Logical Statements
Statements
Identities
Sets
Scores
Lexicon
Objects