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 $\mathbf{E}((Ax + b - y)^\top (Ax + b - y))$

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

The gradients with respect to b are

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

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

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

[1] Future editions will include an account.

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

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

$$A\mathsf{E}(x) + b - \mathsf{E}(y) = 0$$
$$\mathsf{E}(xx^\top)A^\top + \mathsf{E}(x)b^\top - \mathsf{E}(xy^\top) = 0$$

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

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

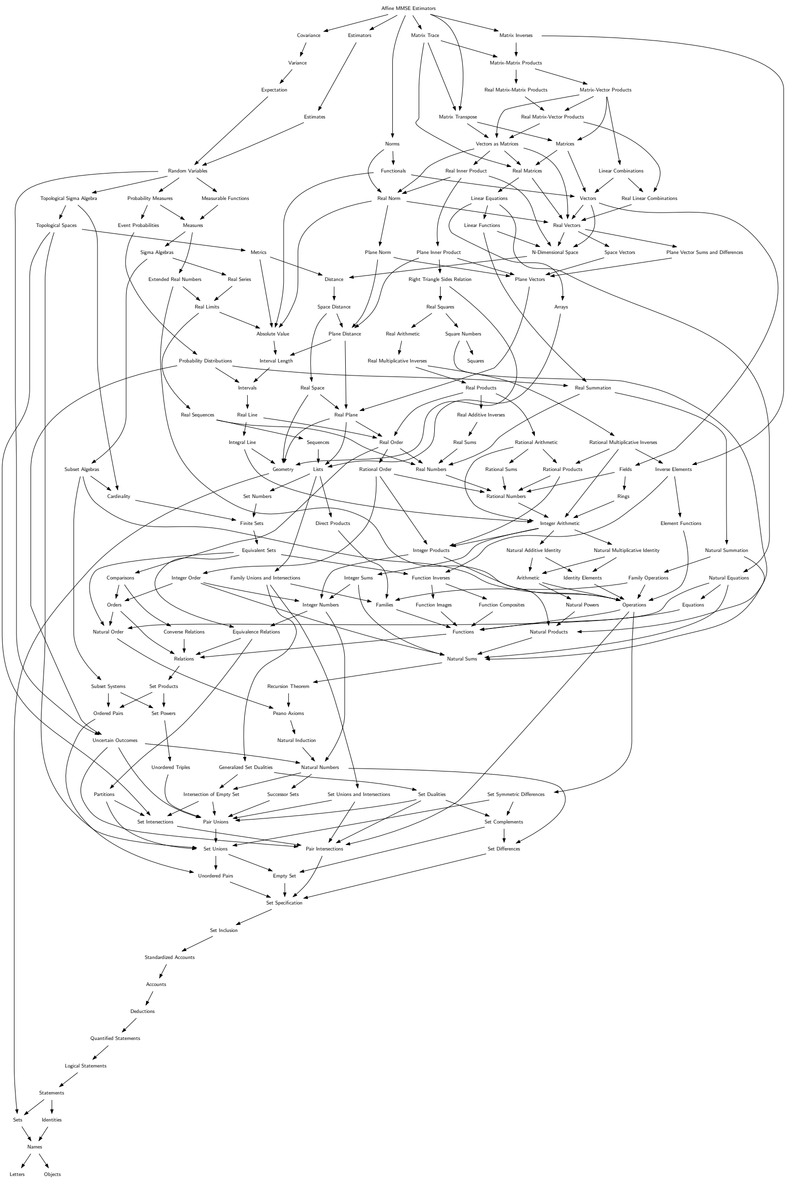
So $A^\top = \operatorname{cov}(x,x)^{-1}\operatorname{cov}(x,y)$ means $A = \operatorname{cov}(y,x)\operatorname{cov}(x,x)^{-1}$ is a solution. Then $b = \mathsf{E}(y) - \operatorname{cov}(y,x)\operatorname{cov}(x,x)^{-1}\mathsf{E}(x)$. So to summarize, the estimator $\phi(x) = Ax + b$ is

$$\operatorname{cov}(y,x)\,(\operatorname{cov} x,x)^{-1}\,x + \mathsf{E}(y) - \operatorname{cov}(y,x)\operatorname{cov}(x,x)^{-1}\mathsf{E}(x)$$

or

$$\mathsf{E}(y) + \operatorname{cov}(y,x)\,(\operatorname{cov} x,x)^{-1}\,(x - \mathsf{E}(x))$$

$\square$



Affine MMSE Estimators
Covariance
Estimators
Matrix Trace
Matrix Inverses
Variance
Matrix-Matrix Products
Expectation
Real Matrix-Matrix Products
Matrix-Vector Products
Estimates
Matrix Transpose
Real Matrix-Vector Products
Norms
Vectors as Matrices
Matrices
Random Variables
Functionals
Real Inner Product
Real Matrices
Linear Combinations
Topological Sigma Algebra
Probability Measures
Measurable Functions
Real Norm
Linear Equations
Vectors
Real Linear Combinations
Topological Spaces
Event Probabilities
Measures
Linear Functions
Real Vectors
Sigma Algebras
Metrics
Plane Norm
Plane Inner Product
N-Dimensional Space
Space Vectors
Plane Vector Sums and Differences
Extended Real Numbers
Real Series
Distance
Right Triangle Sides Relation
Plane Vectors
Real Limits
Space Distance
Real Squares
Arrays
Absolute Value
Plane Distance
Real Arithmetic
Square Numbers
Probability Distributions
Interval Length
Real Multiplicative Inverses
Squares
Intervals
Real Space
Real Products
Real Summation
Real Sequences
Real Line
Real Plane
Real Additive Inverses
Integral Line
Sequences
Real Order
Real Sums
Rational Arithmetic
Rational Multiplicative Inverses
Subset Algebras
Geometry
Lists
Rational Order
Real Numbers
Rational Sums
Rational Products
Fields
Inverse Elements
Cardinality
Set Numbers
Rational Numbers
Rings
Finite Sets
Direct Products
Integer Arithmetic
Element Functions
Equivalent Sets
Integer Products
Natural Additive Identity
Natural Multiplicative Identity
Natural Summation
Comparisons
Integer Order
Family Unions and Intersections
Integer Sums
Function Inverses
Arithmetic
Identity Elements
Family Operations
Natural Equations
Orders
Integer Numbers
Families
Function Images
Function Composites
Natural Powers
Operations
Equations
Natural Order
Converse Relations
Equivalence Relations
Functions
Natural Products
Relations
Natural Sums
Subset Systems
Set Products
Recursion Theorem
Ordered Pairs
Set Powers
Peano Axioms
Uncertain Outcomes
Natural Induction
Unordered Triples
Generalized Set Dualities
Natural Numbers
Partitions
Intersection of Empty Set
Successor Sets
Set Unions and Intersections
Set Dualities
Set Symmetric Differences
Set Intersections
Pair Unions
Set Complements
Set Unions
Pair Intersections
Set Differences
Unordered Pairs
Empty Set
Set Specification
Set Inclusion
Standardized Accounts
Accounts
Deductions
Quantified Statements
Logical Statements
Statements
Sets
Identities
Names
Letters
Objects