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

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

[1] Future editions will include an account.

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

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

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

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

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

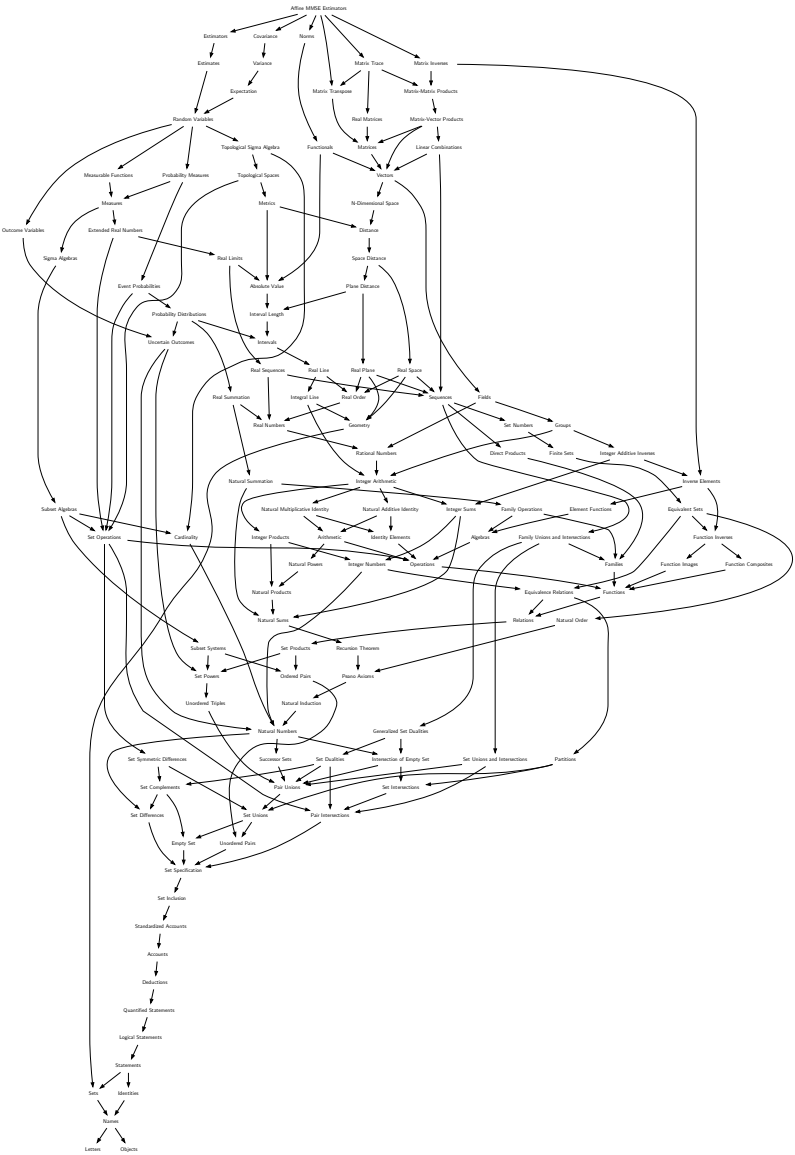
So $A^\top = \operatorname{cov}(x,x)^{-1}\operatorname{cov}(x,y)$ means $A = \operatorname{cov}(y,x)\operatorname{cov}(x,x)^{-1}$ is a solution. Then $b = \mathbf{E}(y) - \operatorname{cov}(y,x)\operatorname{cov}(x,x)^{-1}\mathbf{E}(x)$. So to summarize, the estimator $\phi(x) = Ax + b$ is

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

or

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

$\square$


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