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.[1] 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

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

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

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

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

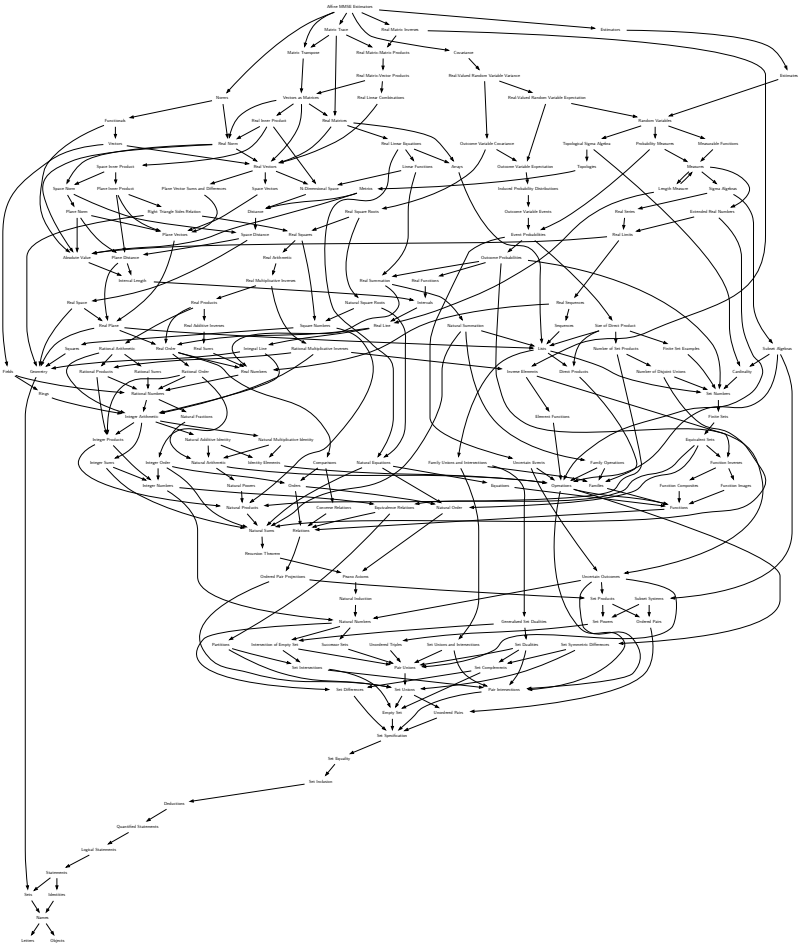
So $A^\top = (\operatorname{cov}(x,x))^{-1}\operatorname{cov}(x,y)$ means $A = \operatorname{cov}(y,x)(\operatorname{cov}(x,x))^{-1}$ is a solution. Then $b = \mathbf{E}(y) - \operatorname{cov}(y,x)\operatorname{cov}(x,x)^{-1}\mathbf{E}(x)$. So to summarize, the estimator $\phi(x) = Ax + b$ is

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

or

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

$\square$


Affine MMSE Estimators
Matrix Trace
Real Matrix Inverses
Estimators
Matrix Transpose
Real Matrix Matrix Products
Concatenated
Real Matrix Vector Products
Real Valued Random Variable Variance
Measures
Vectors
Real Linear Combinations
Real Valued Random Variable Expectation
Functionals
Real Vector Product
Real Matrices
Random Variables
Vectors
Real Sums
Real Linear Equations
Outcome Variable Covariance
Topological Sigma Algebra
Probability Measures
Measurable Functions
Square Norm Product
Real Vectors
Linear Functions
Scalars
Outcome Variable Expectation
Topologies
Measures
Square Norm
Plane Inner Product
Plane Vector Sums and Differences
Square Vectors
N-Dimensional Spaces
Metrics
Natural Probability Distributions
Length Measure
Sigma Algebra
Plane Norm
Right Triangle Sides Relation
Distance
Real Square Roots
Real Series
Extended Real Numbers
Plane Metrics
Natural Distance
Real Squares
Event Probabilities
Real Limits
Absolute Value
Plane Distance
Outcome Probabilities
Real Arithmetic
Natural Length
Real Multiplication Inverses
Real Summation
Real Functions
Real Sequences
Real Spaces
Real Products
Natural Square Roots
Natural
Real Sequences
Real Plane
Real Addition Inverses
Square Numbers
Real Line
Natural Summation
Sum of Series Product
Sequence
Limits
Number of Set Products
Finite Set Enumerable
Subset Algebras
Natural
Natural Arithmetic
Real Group
Real Series
Integral Line
Natural Multiplication Inverses
Finite
Natural Subsets
Natural Order
Real Numbers
Inverse Elements
Quotient Products
Number of Disjoint Unions
Cardinality
Natural Numbers
Set Operators
Integer
Integer Arithmetic
Natural Functions
Element Functions
Finite Sets
Integer Products
Natural Addition Identity
Natural Multiplication Identity
Equivalence Sets
Integer Sums
Integer Order
Natural Arithmetic
Identity Elements
Compositions
Natural Equations
Family Unions and Intersections
Uncertain Events
Family Operations
Function Inverses
Equations
Operators
Functions
Function Images
Integer Numbers
Natural Powers
Order
Function Composites
Natural Products
Concrete Relations
Equivalence Relations
Natural Order
Functions
Natural Sums
Relations
Recursion Theorem
Ordered Pair Projections
Peano Axioms
Uncertain Outcomes
Natural Induction
Set Products
Indexed Systems
Natural Numbers
Generalized Set Dualities
Set Powers
Ordered Pairs
Partitions
Intersection of Empty Set
Incomplete Sets
Unordered Triples
Set Unions and Intersections
Set Dualities
Set Symmetric Differences
Set Intersections
Pair Unions
Set Complements
Set Differences
Set Unions
Pair Intersections
Empty Set
Unordered Pairs
Set Specification
Set Equality
Set Inclusion
Deductions
Quantified Statements
Logical Statements
Statements
Sets
Identities
Terms
Letters
Objects