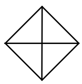


Why

[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

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

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

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

The gradients with respect to b are

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

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

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

[1] Future editions will include an account.

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

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

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

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

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

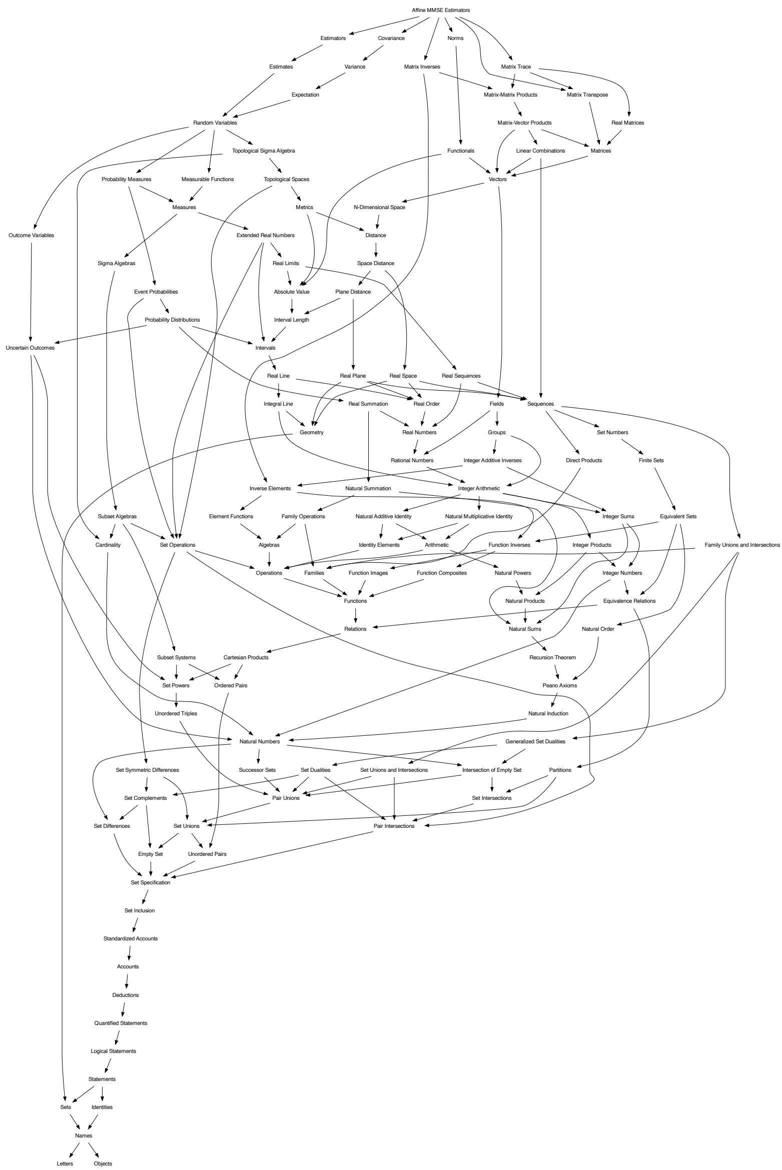
So $A^\top = \mathrm{cov}(x,x)^{-1}\,\mathrm{cov}(x,y)$ means $A = \mathrm{cov}(y,x)\mathrm{cov}(x,x)^{-1}$ is a solution. Then $b = \mathrm{E}(y) - \mathrm{cov}(y,x)\mathrm{cov}(x,x)^{-1}\,\mathrm{E}(x)$. So to summarize, the estimator $\phi(x) = Ax + b$ is

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

or

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

$\square$


Affine MMSE Estimators
Estimators
Covariance
Norms
Estimates
Variance
Matrix Inverses
Matrix Trace
Expectation
Matrix-Matrix Products
Matrix Transpose
Random Variables
Matrix-Vector Products
Real Matrices
Topological Sigma Algebra
Functionals
Linear Combinations
Matrices
Probability Measures
Measurable Functions
Topological Spaces
Vectors
Measures
Metrics
N-Dimensional Space
Outcome Variables
Extended Real Numbers
Distance
Sigma Algebras
Real Limits
Space Distance
Event Probabilities
Absolute Value
Plane Distance
Probability Distributions
Interval Length
Uncertain Outcomes
Intervals
Real Line
Real Plane
Real Space
Real Sequences
Integral Line
Real Summation
Real Order
Fields
Sequences
Geometry
Real Numbers
Groups
Set Numbers
Rational Numbers
Integer Additive Inverses
Direct Products
Finite Sets
Inverse Elements
Natural Summation
Integer Arithmetic
Subset Algebras
Element Functions
Family Operations
Natural Additive Identity
Natural Multiplicative Identity
Integer Sums
Equivalent Sets
Cardinality
Set Operations
Algebras
Identity Elements
Arithmetic
Function Inverses
Integer Products
Family Unions and Intersections
Operations
Families
Function Images
Function Composites
Natural Powers
Integer Numbers
Functions
Natural Products
Equivalence Relations
Relations
Natural Sums
Natural Order
Subset Systems
Cartesian 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
Set Unions and Intersections
Intersection of Empty Set
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
Names
Letters
Objects