Why

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

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

The gradients with respect to b are

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

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

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

---

[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

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

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

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

or

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

$\square$

Affine MMSE Estimation

Covariance — Estimation — Matrix Trace — Matrix Inverse

Variance — Real Matrices — Matrix Transpose — Matrix-Matrix Products

Expectation — Estimates — Matrices

Random Variables — Norms

Measurable Functions — Functionals

Outcome Variables — Topological Sigma Algebra

Topological Spaces — Vectors

Metrics — N-Dimensional Space

Absolute Value — Distance

Probability Measure — Space Distance

Event Probabilities — Measure — Plane Distance

Probability Distributions — Extended Real Numbers

Interval Length

Real Limits — Intervals

Real Summation — Real Sequence — Real Line — Real Plane — Real Space

Real Order — Integral Line — Fields

Real Numbers — Sequences — Geometry — Groups

Natural Summation — Sigma Algebra — Rational Numbers — Direct Products — Integer Additive Inverse — Set Numbers

Integer Arithmetic — Inverse Elements — Finite Sets

Subset Algebra — Cardinality — Natural Additive Identity — Natural Multiplication Identity — Family Operations — Element Functions — Integer Sums — Equivalent Sets

Set Operation — Arithmetic — Identity Elements — Algebra — Integer Products — Family Unions and Intersections — Function Inverse

Natural Exponents — Operations — Families — Integer Number — Function Images — Function Composition

Natural Products — Functions — Equivalence Relations — Natural Order

Uncertain Outcomes — Natural Sums — Relations

Subset Systems — Cartesian Products — Recursion Theorem

Set Power — Ordered Pairs — Peano Axioms

Unordered Triples — Natural Induction

Generalized Set Dualities — Natural Numbers

Set Symmetric Differences — Set Dualities — Set Unions and Intersections — Successor Sets — Intersection of Empty Set — Partitions

Set Complements — Pair Unions — Set Intersections

Set Unions — Set Differences — Pair Intersections

Unordered Pairs — Empty Set

Set Specification

Set Inclusion

Standardized Accounts

Accounts

Deductions

Quantified Statements

Logical Statements

Statements

Identities — Sets

Names

Letters — Objects