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

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

The gradients with respect to b are

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

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

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

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

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

or

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

$\square$

Affine MMSE Estimation

Estimation • Real Matrix Inverses • Matrix Trace

Estimators • Covariance • Real Matrix-Matrix Products • Matrix Transpose

Variance • Real Matrix-Vector Products

Expectation • Real Linear Combinations • Vectors as Matrices • Norms • Functionals

Random Variables • Real Matrices • Real Inner Product • Real Norm • Vectors

Measurable Functions • Probability Measures • Topological Sigma Algebra • Outcome Variable Expectation • Linear Equations • Real Vectors • Square Inner Product

Measures • Topologies • Outcome Variable Covariance • Arrays • Linear Functions • Metrics • Space Vectors • N-Dimensional Space • Plane Vector Sums and Differences • Plane Inner Product • Square Norm

Sigma Algebras • Induced Probability Distributions • Distance • Right Triangle Sides Relation • Plane Norm

Extended Real Numbers • Real Series • Outcome Variable Events • Real Squares • Plane Vectors • Square Distance

Real Limits • Plane Distance • Absolute Value

Event Probabilities

Probability Distributions • Real Arithmetic

Outcome Probabilities • Real Multiplication Inverses • Interval Length

Real Summation • Natural Square Roots • Intervals • Real Products • Real Plane

Square Numbers • Real Line • Real Addition Inverses • Real Space

Integral Line • Real Sums • Real Order • Rational Arithmetic • Rational Multiplication Inverses • Square • Geometry • Fields

Subset Algebras • Natural Summation • Sequences • Inverse Elements • Real Numbers • Rational Order • Rational Sums • Rational Products • Rings

Cardinality • Lists • Direct Products • Rational Numbers

Set Numbers • Element Functions • Finite Sets • Equivalent Sets • Natural Multiplication Identity • Natural Addition Identity • Integer Products

Uncertain Events • Operations • Family Operations • Identity Elements • Arithmetic • Natural Powers • Integer Sums

Function Inverses • Family Unions and Intersection • Natural Equations • Comparators • Orders • Integer Order • Natural Products • Integer Numbers

Function Image • Function Composition • Families • Equations • Natural Order • Equivalence Relations • Converse Relation • Natural Sums

Functions • Relations • Recursion Theorems

Uncertain Outcomes • Peano Axioms • Ordered Pair Projections

Subset Systems • Set Products • Natural Induction

Ordered Pairs • Set Powers • Generalized Set Duality • Natural Numbers • Partitions

Set Symmetric Differences • Set Duality • Set Unions and Intersection • Unordered Triples • Intersection of Empty Set • Successor Sets

Set Complements • Pair Unions • Set Intersections

Pair Intersections • Set Differences • Set Unions

Unordered Pairs • Empty Set

Set Specification

Set Equality

Set Inclusion

Deductions

Quantified Statements

Logical Statements

Statements

Identities • Sets

Letters • Objects