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

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

The gradients with respect to $b$ are

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

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

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

---

[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

$$AE(x) + b - \mathbf{E}(y) = 0$$

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

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

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

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 Estimators

Estimators  Covariance  Norms  Matrix Inverse  Matrix Trace

Estimates  Variance  Matrix-Matrix Products  Matrix Transpose

Expectation  Matrix-Vector Products  Real Matrices

Random Variables  Functionals  Linear Combinations  Matrices

Topological Sigma Algebra

Probability Measure  Measurable Functions  Topological Spaces  Vectors

Measure  Metrics  N-Dimensional Space

Outcome Variables  Sigma Algebras  Extended Real Numbers  Distance

Event Probabilities  Real Limits  Space Distance

Probability Distributions  Absolute Value  Plane Distance

Uncertain Outcomes  Interval Length

Intervals

Real Line  Real Plane  Real Space  Real Sequences

Integer Line  Real Summation  Real Order  Fields  Sequences

Geometry  Real Numbers  Groups  Direct Products  Set Numbers

Rational Numbers  Integer Additive Inverse  Finite Sets

Inverse Elements  Natural Summation  Integer Arithmetic  Integer Sums

Subset Algebra  Element Functions  Family Operations  Natural Additive Identity  Natural Multiplicative Identity  Function Inverses  Equivalent Sets

Cardinality  Set Operations  Algebras  Identity Elements  Arithmetic  Function Inverse  Integer Products  Family Unions and Intersections

Operations  Families  Function Image  Function Composites  Natural Powers  Integer Numbers

Functions  Natural Products  Equivalence Relations

Subset Systems  Cartesian Products  Relations  Natural Sums  Natural Order

Set Powers  Ordered Pairs  Recursion Theorem

Unordered Triples  Peano Axioms

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

Empty Set  Unordered Pairs  Pair Intersections

Set Specification

Set Inclusion

Standardized Accounts

Accounts

Deductions

Quantified Statements

Logical Statements

Statements

Sets  Identities

Names

Letters  Objects

4