Affine MMSE Predictors

We want to find A and b to minimize

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

Proof. We can 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

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

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

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

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

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

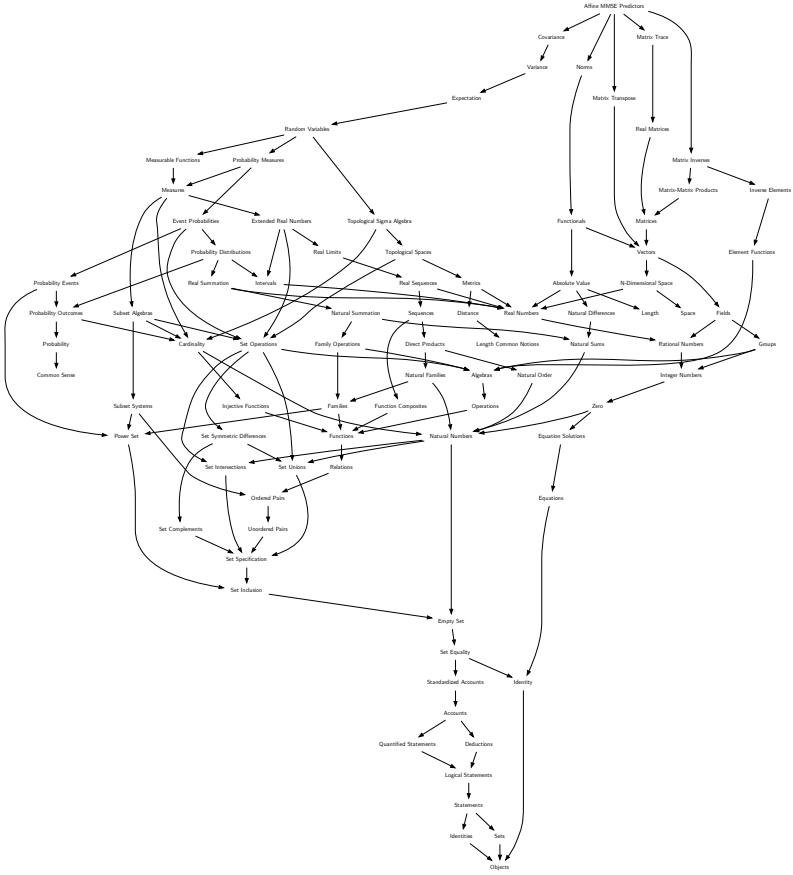
So $A^\top = \mathbf{cov}(x,x)^{-1}\,\mathbf{cov}(x,y)$ means $A = \mathbf{cov}(y,x)\,\mathbf{cov}(x,x)^{-1}$ is a solution. Then $b = \mathbf{E}(y) - \mathbf{cov}(y,x)\,\mathbf{cov}(x,x)^{-1}\,\mathbf{E}(x)$. So to summarize, the estimator $\phi(x) = Ax + b$ is

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

or

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

$\square$


Affine MMSE Predictors
Covariance
Matrix Trace
Variance
Norms
Expectation
Matrix Transpose
Random Variables
Real Matrices
Measurable Functions
Probability Measures
Matrix Inverses
Matrix-Matrix Products
Inverse Elements
Measures
Event Probabilities
Extended Real Numbers
Topological Sigma Algebra
Functionals
Matrices
Vectors
Probability Distributions
Real Limits
Topological Spaces
Element Functions
Probability Events
Real Summation
Intervals
Real Sequences
Metrics
Absolute Value
N-Dimensional Space
Probability Outcomes
Subset Algebras
Natural Summation
Sequences
Distance
Real Numbers
Natural Differences
Length
Space
Fields
Probability
Cardinality
Set Operations
Family Operations
Direct Products
Length Common Notions
Natural Sums
Rational Numbers
Groups
Common Sense
Natural Families
Algebras
Natural Order
Integer Numbers
Subset Systems
Injective Functions
Families
Function Composites
Operations
Zero
Power Set
Set Symmetric Differences
Functions
Natural Numbers
Equation Solutions
Set Intersections
Set Unions
Relations
Ordered Pairs
Equations
Set Complements
Unordered Pairs
Set Specification
Set Inclusion
Empty Set
Set Equality
Standardized Accounts
Identity
Accounts
Quantified Statements
Deductions
Logical Statements
Statements
Identities
Sets
Objects