James-Stein Estimator

Situation

You want to estimate unknown parameter W of dimension p >=3 from a single sample X such that the risk of the estimate is as low as possible.

You know the parameter W is drawn from a Gaussian distribution with mean 0 and variance sigma^2:
W ~ N(0,sigma^2 I).

The risk of the estimate is defined as the expected value of the loss.

The loss is defined as the L2 norm of the difference between the unknown parameter W and the estimate you provide e.

Thus risk(W, e) = E Loss(W, e) = E ||W - e||^2.

An Incorrect Answer

Since you have just one sample X, use it as the estimate of W. To make things concrete, suppose that p = 3 and X = [0.4, -.05, 0.18].

The risk is then
risk(W, X) = E Loss(W,X) = E ||W - X||^2 = E ||[w1 - 0.4, w2 + 0.05, w3 - 0.18] || ^ 2.

The Correct Answer

Instead of using X, use Y = (1 - (p - 2) sigma^2/||X||^2) X, which uses the L2 norm of X.

For example if sigma^2 = 1 (so that W ~ N(0,1)), p = 3, then using X as defined above,
Y = (1 - (3 - 2) * 1/||X||^2) X = (1 - 1 / 0.44) X ~= (1 - 2.27) X = -1.27 X.

In james-61, a proof is provided that remarkably
risk(W, Y) < risk(W, X) for all W, provided that p >=3 and W ~ N(0,sigma^2 I).

Continuing the example, consider the point W = 0. Then
risk(0, X) = E || [0.4, -.05, 0.18]||^2 ~= E (0.44) = 0.44
and
Y = (1 - 1 / ||X||^2) ~= (1 - 1 / 0.44) X = (1 - 2.27) X = -1.27 X
risk(0, Y) ~= E || (-1.27) [0.4, -.05, 0.18] ||^2 ~= E (-1.27)(0.44) = E (-0.56) < E (0.44) = risk(0,X).

Another example. Suppose the same was X = [1,2,3] and again W = 0. Then
risk(0,X) = E ||X||^2 = E || [1,2,3] ||^2 ~= E (3.87) = 3.87
and
Y = (1 - 1 / ||X||^2) X ~= (1 - 1 / 3.87) X = 0.74 X
risk(0,Y) ~= E || 0.74 X ||^2 < E ||X||^2.

These examples show that the norm of X must be sufficiently large or the multiplier becomes negative.

Ref: [james-61]. W. James and Charles Stein, “Estimation With Quadratic Loss”, Proc. Fourth Berkeley Symp. on Math Statist and Prob., 1961, pp 361-379. Can be downloaded from http://projecteuclid.org/DPubS?service=UI&version=1.0&verb=Display&handle=euclid.bsmsp/1200512173