Risk Inflation relative to Bayes Oracle

Charles Zheng and Yuval Benjamini

October 21, 2015

These are preliminary notes.

1 Ridge regression

Conjecture:

$$\sup \frac{R_{\lambda^*}(H, \alpha^2, \gamma)}{R_0(H, \alpha^2, \gamma)} < \infty$$

The risk inflation from not knowing λ^* is bounded.

FALSE!! Unbounded...

Identity case.

$$m_I(-\lambda) = \frac{-(1-\gamma+\lambda) + \sqrt{(1-\gamma+\lambda)^2 + 4\gamma\lambda}}{2\gamma\lambda}$$

$$\begin{aligned} m_I'(z) &= \frac{d}{dz}\frac{(1-\gamma-z) - \sqrt{(1-\gamma-z)^2 - 4\gamma z}}{2\gamma z} \\ &= \frac{-1}{z} m_I(z) + \frac{1}{2\gamma z}\left[-1 - \frac{z-\gamma-1}{\sqrt{(z+\gamma-1)^2 - 4\gamma z}}\right] \end{aligned}$$

where

$$\lim_{z\to 0} -1 - \frac{z-\gamma-1}{\sqrt{(z+\gamma-1)^2 - 4\gamma z}} = \frac{2}{\gamma - 1}$$

and

$$
\begin{aligned}
\lim_{z\to 0} m_I(z) &= -\lim_{z\to 0} \frac{(z+\gamma-1)+\sqrt{(z+\gamma-1)^2-4\gamma z}}{2\gamma z} \\
&= -\frac{2(\gamma-1)}{\gamma}\frac{1}{z}
\end{aligned}
$$

Hence,

$$\lim_{\lambda\to 0}(\gamma-\lambda\alpha^2)\lambda m'(-\lambda) = \gamma\left[m(0)-\frac{1}{2\gamma}\frac{2}{\gamma-1}\right]$$

$$
\begin{aligned}
R_0 &= 1-\gamma m(0)+\gamma\lambda m'(0) \\
&= 1-\gamma m(0)+\gamma m(0)-\frac{1}{\gamma-1} = \frac{\gamma-2}{\gamma-1}
\end{aligned}
$$

2 Covariance estimation

$$S \sim W_n(\frac{1}{n}\Sigma), D = \mathrm{diag}(S), \hat{R} = D^{-1/2}SD^{-1/2}$$

$$S_\lambda = \lambda D + (1-\lambda)S$$

Which λ minimizes

$$\mathbf{E}\mathrm{tr}[S_\lambda^{-1}\Sigma] + \log\det(S_\lambda)$$

We have

$$\log\det(\lambda D+(1-\lambda)S) = \log\det D+\log\det(\lambda I+(1-\lambda)\hat{R}) = \log\det D+\sum_{i=1}^{p}\log(\lambda+(1-\lambda)r_i)$$

where r_i are the eigenvalues of $\hat{R}$.

Meanwhile

$$
\begin{aligned}
\mathbf{E}\mathrm{tr}[S_\lambda^{-1}\Sigma] &= \mathbf{E}\mathrm{tr}[(\lambda D+(1-\lambda)S)^{-1}\Sigma] \\
&= \frac{1}{1-\lambda}\mathbf{E}\mathrm{tr}[(\frac{\lambda}{1-\lambda}I+\hat{R})^{-1}D^{-1/2}\Sigma D^{-1/2}]
\end{aligned}
$$

Take $n, p \to \infty$. Then $D^{-1/2}\Sigma D^{-1/2} \to R$, the true correlation. From now we can just assume $\Sigma = R$, (ie unit marginal variances), it doesn't matter in the limit. Then we get

$$\mathbf{E}\mathrm{tr}[S_\lambda^{-1}\Sigma] = \frac{1}{1-\lambda}\mathbf{E}\mathrm{tr}[(\frac{\lambda}{1-\lambda}I + S)^{-1}\Sigma]$$

We know how to evaluate the term inside, i.e. $\mathbf{E}\mathrm{tr}[(\frac{\lambda}{1-\lambda}I + S)^{-1}\Sigma]$ based on random matrix theory.