Discussion of

Proper Bayes Minimax Multiple Shrinkage Estimation

by Ed George (with P. Bhagwat and B. Strawderman)

THE OHIO STATE UNIVERSITY

Christopher Hans

Department of Statistics

15th Objective Bayes Methodology Conference
Athens, Greece

June 10, 2025

Recap

Ed, Pankaj, and Bill found conditions under which there exist proper Bayes, minimax, multiple shrinkage estimators in the normal means problem.

This complements George (1986), who showed that certain multiple shrinkage estimators under improper priors are minimax.

The key is breaking down the sample space into two overlapping regions and carefully studying the behavior of the underlying marginal on the two regions.

A clever *rescaling* of the prior based on the marginal's behavior over the more difficult of the two regions provides conditions under which a proper Bayes estimator that is minimax exists.

Congratulations on an interesting result—it's inspiring to see long-term dedication to thinking about a problem pay off!

Strawderman prior

A key example prior is the proper Strawderman prior:

$$X \mid \mu \sim \mathrm{N}\left(\mu, I_p\right)$$
$$\mu \mid \lambda \sim \mathrm{N}\left(0, \frac{1-\lambda}{\lambda} I_p\right)$$
$$\lambda \sim \lambda^{-\alpha}/(1-\alpha)$$

In the context of shrinkage toward multiple targets, $\theta_1, \ldots, \theta_K$, a minimax estimator of μ can be constructed by replacing the marginal density $m_\pi(t)$ and its corresponding $r(t)$ with $m_a(t) \propto m_\pi(t / a)$ and $r_a(t) = r(t / a)$, where $a > 0$ is an appropriately chosen constant.

- For some α and $\theta_1, \ldots, \theta_K$, $a = 1$ suffices and the estimator under the proper Strawderman prior is minimax.
- Otherwise, the resulting estimator can be thought of as arising from the prior on λ that generates the marginal density $m_a(t)$.

Rescaling

Appropriate scaling factors *a* satisfy:

$$a \geq D \boxed{\left(1 + \frac{1}{\rho}\right)^2 \frac{1}{\sqrt{r^{-1}(p-2)}}}$$

$$D = \max_{i \neq j} \|\theta_i - \theta_j\|^2$$

Depends only on properties of $r(t)$, or, equivalently, $m_\pi(t)$.

For a given class of priors, the choice of shrinkage targets drives the rescaling.

What happens in the extreme cases?

For what values of a and D is the multiple shrinkage estimator under the proper Strawderman prior minimax?

Shrinkage Targets

What happens in the extreme cases?

When the shrinkage targets are all close to each other, $D \approx 0$, and $a = 1$ (no rescaling) should suffice.

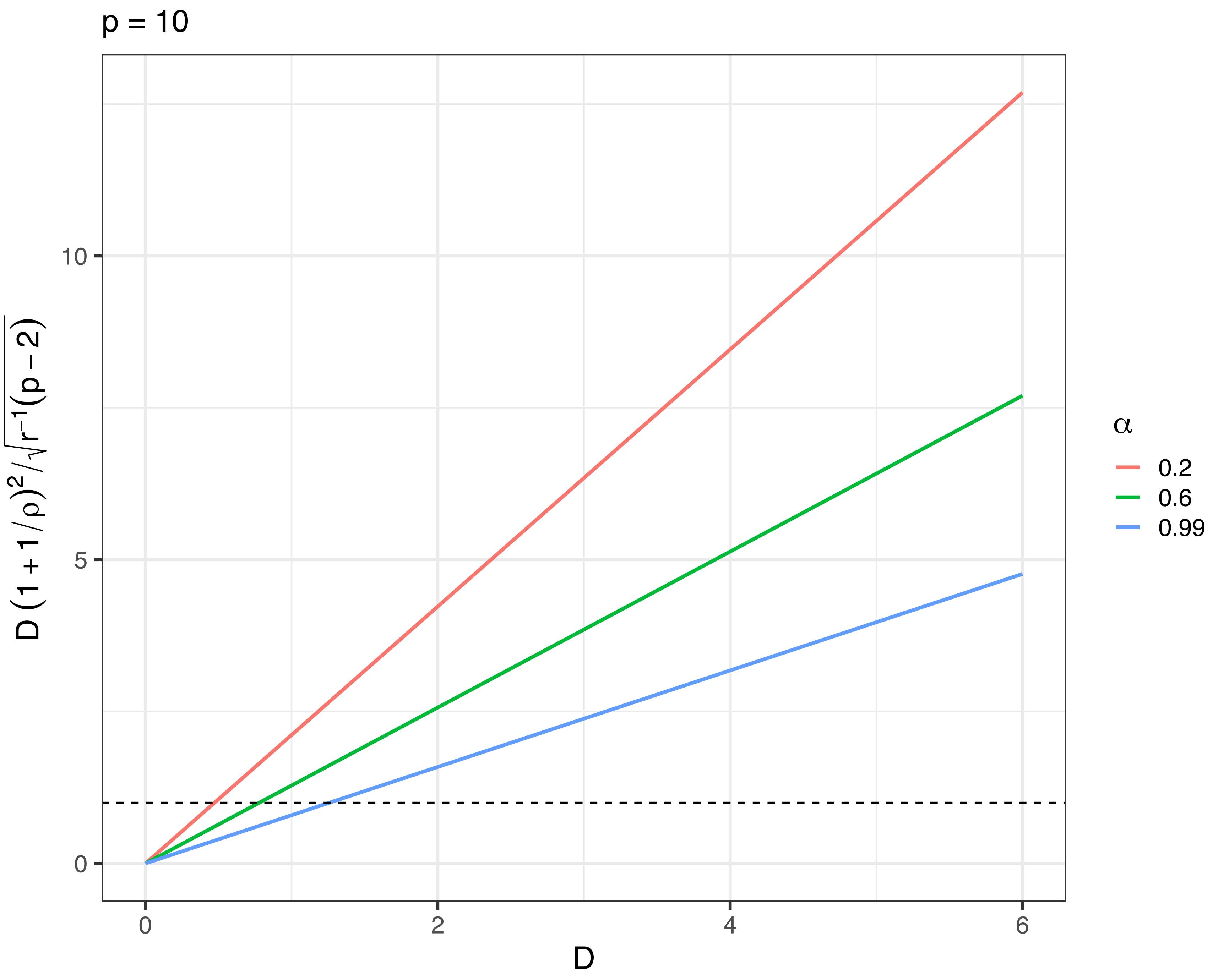
Makes sense: aren't really doing "multiple" shrinkage and the usual minimax result for shrinkage toward a single target should apply.

When at least one shrinkage target is far apart from one of the others, D will be large, and a large amount of rescaling ($a \gg 1$) is needed.

$r(t_i)$ is increasing for Strawderman's prior, so large a results in less shrinkage.

Shrinkage Targets

For what values of α and D is the multiple shrinkage estimator under the proper Strawderman prior minimax?



Global vs. Local Rescaling

The amount of rescaling required is driven by the pair of shrinkage targets that are farthest away from each other.

The scaling applies globally to all components of X.

Rich literature on global vs. local shrinkage (e.g., Carvalho, Polson and Scott, 2010; Polson and Scott, 2010; Polson and Scott, 2012; van der Pas, *et al.*, 2014, 2017; *etc.*).

Som, Hans and MacEachern (2014, 2016) describe undesirable consequences of global mono-shrinkage in regression (OBayes 2015).

- Could the components of X be rescaled locally so that the rescaling has less of an impact?

Shrinkage Targets and Weights

The amount of rescaling required to guarantee minimaxity depends on both the shrinkage targets, $\theta_1, \ldots, \theta_K$, and the properties of the underlying marginal distribution, $m_\pi(x)$, but not on the shrinkage target weights, w_i.

Do we need to be concerned about one shrinkage target, θ_i, that is far away from the others if the corresponding weight, w_i, is very small?

In weak prior information settings, do the results help us think about how to select shrinkage targets and/or corresponding weights?

Of interest to me: Hans, Peruggia and Wang (2023) examine the interplay between shrinkage targets and marginal likelihoods in regression with influential observations.

Conclusion

Thanks for a great talk!