Electronic Journal of Statistics
Vol. 9 (2015) 811–822
ISSN: 1935-7524
DOI: 10.1214/15-EJS1000REJ

Rejoinder of "High-dimensional autocovariance matrices and optimal linear prediction"*

Timothy L. McMurry

Department of Public Health Sciences
University of Virginia
Charlottesville, VA 22908
e-mail: tmcmurry@virginia.edu

and

Dimitris N. Politis

Department of Mathematics
University of California, San Diego
La Jolla, CA 92093
e-mail: dpolitis@ucsd.edu

Received March 2015.

We would like to sincerely thank all discussants for their kind remarks and insightful comments. To start with, we wholeheartedly welcome the proposal of Rob Hyndman for a "better acf" plot based on our vector estimator $\hat{\gamma}^*(n)$ from Section 3.2. As mentioned, the sample autocovariance is not a good estimate for the vector $\gamma(n)$, and this is especially apparent in the wild excursions it takes at higher lags—see the left panel of Figure 1 of Hyndman's piece. Note that these wild (and potentially confusing) excursions are the norm rather than the exception; they are partly explainable by two facts: (a) the identity $\sum_{k=-n}^{n} \breve{\gamma}_k = 0$ implies that $\breve{\gamma}_k$ must misbehave for higher lags to counteract its good behavior for small lags; and (b) the $\breve{\gamma}_k$ are correlated, and therefore their excursions appear smooth (and may be confused for structure). The only saving point of the current `acf` plot in R is that it has a `lag.max` default of $10 \log_{10} n$ so the ugliness occuring at higher lags is masked. Interestingly, showing just the lags up to $10 \log_{10} n$ is tantamount to employing a rectangular lag-window—which is one of the flat-top kernels albeit not the best—with a logarithmic choice for l that is indeed optimal under the exponential decay of γ_k typical of ARMA models.

Rob Hyndman also brings up the question of optimal linear prediction. Here, we would just like to offer a linguistic comment. The statistical term 'optimal estimation' is clear: an optimal estimator is closest (according to some criterion) to its target estimand. However, the term 'optimal prediction' is typically used in a probabilistic context where all parameters are assumed known and only the form of the predictor is in question; in other words, the term 'optimal prediction'

*Main article 10.1214/15-EJS1000.