Optimality of the Karhunen-Loeve basis in nonlinear reconstruction

Stephane Mallat,[1] and Ofer Zeitouni [2]

September 4, 2000.

General problem Let $\mathbf{Y} := (Y_1, \ldots, Y_N)$ denote an N-dimensional Gaussian vector with independent zero mean components of variance $\sigma_i = E(Y_i^2)$. We assume for concreteness that $\sigma_i \geq \sigma_{i+1}$.

Let $T(\theta)$ denote an arbitrary orthogonal matrix on $\mathbb{R}^N$ (θ is a $N(N-1)/2$ dimensional parameter, and we take $T(0) = I$), and define the random variable $\mathbf{X}(\theta) = T(\theta)\mathbf{Y}$. For any $M < N$, define

$$\mathcal{E}(N, M, \theta) = E\Big(\min_{i_1 \neq i_2 \neq \ldots \neq i_{N-M}} \sum_{j=1}^{N-M} (X_{i_j}(\theta))^2 \Big) .$$

$\mathcal{E}(N, M, \theta)$ is the mean square error when reconstructing $\mathbf{X}(\theta)$ according to its M largest (in absolute value) components. When the matrix $T(\theta)$ is a permutation matrix, this is the reconstruction error when keeping the M largest (in absolute value) components of $\mathbf{Y}$, where $\mathbf{Y}$ is already expressed in its Karhunen-Loeve basis, whereas other choices of θ correspond to an expansion in other, non K-L bases.

Let $T(\theta) \in \mathcal{P}$ if $T(\theta)$ is composed only of zeroes and ones, i.e. $T(\theta)$ is a permutation and reflection matrix. We have the following

Conjecture 1.

$$\min_{\theta} \mathcal{E}(N, M, \theta) = \min_{T(\theta) \in \mathcal{P}} \mathcal{E}(N, M, \theta) = \mathcal{E}(N, M, 0) .$$

Conjecture 1, if true, implies that the Karhunen-Loeve basis is the best basis not only for linear reconstruction but also for nonlinear reconstruction based on the M largest projections.

We do not know how to prove in general Conjecture 1. However, it does hold true for $M = 1$, i.e reconstruction based on the largest projection. Indeed, we have

Lemma 1. *Conjecture 1 holds true if $M = 1$.*

Proof: We can re-parametrize T such that $\eta_i = E(X_i^2)$ satisfies $\eta_i \geq \eta_{i+1}$. Note that $\eta_i = \sum_{j=1}^{N} T_{ij}^2 \sigma_j$. Clearly, it is enough then to prove that $E(\max X_i^2) \leq E(\max Y_i^2)$.

[1]Ecole Polytechnique, Palaiseau
[2]Technion, Haifa

Let $\tilde{\mathbf{X}}$ denote a vector of independent Gaussian random variables with $E\tilde{X}_i = 0$ and $E(\tilde{X}_i)^2 = \eta_i$. By Sidak's inequality [1], for any $t > 0$,

$$P(\cap\{|X_i| < t\}) \geq \prod P(|X_i| < t) = P(\cap\{|\tilde{X}_i| < t\}),$$

implying that

$$E(\max|X_i|^2) \leq E(\max|\tilde{X}_i|^2). \tag{1}$$

Next, we can check that

$$\sum_{i=1}^{m} \sigma_i \geq \sum_{i=1}^{m} \eta_i,\ i = 1, \ldots, N-1. \tag{2}$$

and clearly, because $\sum_i T_{ij}^2 = 1$, also

$$\sum_{i=1}^{N} \sigma_i = \sum_{i=1}^{N} \eta_i.$$

Indeed, when $m = 1$, (2) holds because $\sum_j T_{1j}^2 = 1$. For $m = 2$, we have that

$$\begin{aligned}
\eta_1 + \eta_2 &= \sum_j (T_{1j}^2 + T_{2j}^2)\sigma_j \\
&= \sum_{j=2}^{N} (T_{1j}^2 + T_{2j}^2)\sigma_j + (T_{11}^2 + T_{21}^2)\sigma_1 \\
&\leq \sigma_2 \sum_{j=2}^{N} (T_{1j}^2 + T_{2j}^2) + (T_{11}^2 + T_{21}^2)\sigma_1 \\
&= \sigma_2(2 - T_{11}^2 - T_{21}^2) + (T_{11}^2 + T_{21}^2)\sigma_1 \\
&= \sigma_1 + \sigma_2 + (\sigma_2 - \sigma_1)(1 - T_{11}^2 - T_{21}^2) \leq \sigma_1 + \sigma_2.
\end{aligned}$$

The general case of (2) follows by induction.

By an inequality of Marshall and Proschan, see [2, Application 7.A.18], one concludes that for any convex, permutation symmetric function ϕ,

$$E(\phi(\tilde{X}_1^2, \ldots, \tilde{X}_N^2)) \leq E(\phi(Y_1^2, \ldots, Y_N^2)).$$

Applying this to the function $\phi(\cdot) = \max x_i^2$, one concludes that $E(\max Y_i^2) \geq E(\max(\tilde{X}_i)^2)$, which together with (1) yields $E(\max Y_i^2) \geq E(\max X_i^2)$, as claimed. $\square$

Remark: the Schur convexity part of the argument holds also for the function

$$\phi_M(\mathbf{x}) = \max_{i_1 \neq i_2 \neq \ldots \neq i_M} \sum_{j=1}^{M} x_{i_j}^2.$$

What is missing in order to prove the conjecture for general M is the analog of (1): is it true that

$$E(\phi_M(\mathbf{X})) \leq E(\phi_M(\tilde{\mathbf{X}}))\,?$$

References

[1] Das Gupta S., Eaton M. L., Olkin I., Perlman M., Savage L. J., Sobel M., Inequalities and the probability content of convex regions for elliptically contourned distributions, Proc. sixth Berkeley Symp. Math. Stat. Prob., (1972), 241–264.

[2] M. Shaked and J. G. Shantikumar, Stochastic orders and their applications, Academic Press 1994