A CORRELATION INEQUALITY

FOR NONLINEAR RECONSTRUCTION

Ofer Zeitouni
University of Minnesota and Technion
www.math.umn.edu/$\sim$zeitouni

ENS, JUNE 1 2006

(Based on joint work with S. Mallat)

Data $\mathbf{X} := (X_1, \ldots, X_N)^*$ - Gaussian vector, zero mean, covariance $\mathbf{R}$.

$T \in \mathbf{O}_N$ - fixed orthogonal matrix acting on $\mathbb{R}^N$, columns T_i, $i = 1, \ldots, N$.

$k < N$ - fixed integer.

$a_i = (T\mathbf{X})_i$ - projections onto basis determined by T.

Linear reconstruction:

$$\mathbf{X}^{L,k,T} = \sum_{i=1}^{k} a_i T_i \,.$$

Error: $\mathcal{E}^{L,k,T} = E||\mathbf{X} - \mathbf{X}^{L,k,T}||^2$.

$(\mathbf{X} := (X_1, \dots, X_N)^*,\ a_i = (T\mathbf{X})_i,\ \mathbf{X}^{L,k,T} = \sum_{i=1}^{k} a_i T_i,\ \mathcal{E}^{L,k,T} = E||\mathbf{X} - \mathbf{X}^{L,k,T}||^2)$

Fact: (easy)

If $R = U\Lambda U^*$ with $U \in \mathbf{O}_N$ and Λ diagonal with $\lambda_i = \Lambda_{ii}$ decreasing in i then

$$\min_{T \in \mathbf{O}_N} \mathcal{E}^{L,K,T} = \mathcal{E}^{L,k,U^*} = \sum_{i=k+1}^{N} \lambda_i \,.$$

Proof: Write $V = TU$. Then, can check that

$$E||\mathbf{X} - \mathbf{X}^{L,k,T}||^2 = \sum_{i=k+1}^{N} \sum_{\ell} \lambda_\ell V_{i\ell}^2 = \text{Trace}(V\Lambda V^* D_k)\,,$$

where D_k is diagonal and

$$D_k = \begin{pmatrix} \mathbf{0}_k & \\ & \mathbf{I}_{N-k} \end{pmatrix}.$$

Since this is a convex functional of V and the extreme points of the polytope of doubly stochastic matrices are permutations, the minimum is achieved at a permutation, which is easily checked to be the identity (since the entries of Λ are ordered).

Note that this computation did not depend on the Gaussian structure, only on L^2 arguments.

This solution can be reformulated as follows:
if $\mathbf{Y} := (Y_1, \ldots, Y_N)^*$ has uncorrelated entries,
and $E\mathbf{Y}_i^2 = \lambda_i$, and $\mathbf{X} = U\mathbf{Y}$, then optimal $T = U^*$.

Nonlinear reconstruction: Fix T non-random. Order the a_i according to their absolute value: (i) denotes the random permutation such that

$$|a_{(1)}| \geq |a_{(2)}| \geq \ldots \geq |a_{(N)}| .$$

Set

$$\mathbf{X}^{NL,k,T} = \sum_{i=1}^{k} a_{(i)} T_{(i)},$$

$$\mathcal{E}^{NL,k,T} = E||\mathbf{X} - \mathbf{X}^{NL,k,T}||^2.$$

Question: What is best T now?

Conjecture: For gaussian $\mathbf{X} = U\mathbf{Y}$ with $\mathbf{Y}$ having independent entries and $E\mathbf{Y}_i^2 = \lambda_i$, optimal T is still U^*.

Natural: "What else could it be"?

Nonlinearity of reconstruction prevents direct use of L^2 computations.

Reformulate: $\mathbf{Y} := (Y_1, \ldots, Y_N)^*$ Gaussian, zero mean, independent entries, variances $\lambda_i \geq \lambda_{i+1}$. $\mathbf{X} = T\mathbf{Y}$.

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

Conjecture: $\min_T \mathcal{E}(N,k,T) = \mathcal{E}(N,k,I_N)$.

$(\mathbf{Y} := (Y_1, \ldots, Y_N)^*$, $\mathbf{X} = T\mathbf{Y}$, $\mathcal{E}(N,k,T) = E(\min_{i_1 \neq i_2 \neq \ldots \neq i_{N-k}} \sum_{j=1}^{N-k} (X_{i_j})^2))$

Conjecture true if k=1 (Mallat-Z.)

Consider first $k = 1, N = 2$. Set $\mathbf{Z} = (Z_1, Z_2)$, Z_i-Gaussian, zero mean, unit variance, $\rho = EZ_1Z_2$.

$$\sigma_+^2 = (1-\alpha)\sigma_1^2 + \alpha\sigma_2^2\,, \qquad \sigma_-^2 = \alpha\sigma_1^2 + (1-\alpha)\sigma_2^2\,.$$

$$h(Z_1, Z_2) = \sigma_+^2 Z_1^2 \mathbf{1}_{\sigma_+^2 Z_1^2 < \sigma_-^2 Z_2^2} + \sigma_-^2 Z_2^2 \mathbf{1}_{\sigma_+^2 Z_1^2 > \sigma_-^2 Z_2^2}$$

$$I(\alpha, \rho) = Eh(Z_1, Z_2)\,.$$

Claims (compute!): 1) $\frac{\partial I(\alpha,\rho)}{\partial \rho} \geq 0\,, \quad \rho > 0$

2) $I(\alpha, \rho)$ is concave in α.

$\Rightarrow$ Conjecture for $N = 2, k = 1$ since $\alpha = 0, \rho = 0$ or $\alpha = 1, \rho = 0$ optimal.

General N with $k = 1$:

$\mathbf{Y} := (Y_1, \ldots, Y_N)^*$, $\mathbf{X} = T\mathbf{Y}$,

$$\mathcal{E}(N, 1, T) = E(\min_{i_1 \neq i_2 \neq \ldots \neq i_{N-k}} \sum_{j=1}^{N-k} (X_{i_j})^2)) = E|\mathbf{Y}|_2^2 - E(\max_{i=1}^{N} X_i^2)$$

Want: $E(\max_i X_i^2)$ is maximal when $T = I_N$.

Need two correlation inequalities.

1) (Sidak): Let μ_N be a centered Gaussian measure on $\mathbb{R}^N$. Let $A \subset \mathbb{R}^N$ be convex and symmetric ($x \in A \Rightarrow -x \in A$). Let $B = \{x \in \mathbb{R}^N : |x_1| < 1\}$. Then,

$$\mu_N(A \cap B) \geq \mu_N(A)\mu_N(B)\,. \tag{1}$$

`Remark:` There is a long standing conjecture that (1) is true as soon as also B is convex symmetric. Best available result is by Hargé, for B an ellipsoid.

2) (Schur convexity) Let $\mathbf{Z} = (Z_1, \ldots, Z_N)$, $\tilde{\mathbf{Z}} = (\tilde{Z}_1, \ldots, \tilde{Z}_N)$ be vectors of independent centered Gaussians, with $EZ_i^2 = \sigma_i^2$ and $E\tilde{Z}_i^2 = \eta_i^2$. Assume that

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

$$\sum_{i=1}^{N} \sigma_i^2 = \sum_{i=1}^{N} \eta_i^2 .$$

Then, for any convex function ϕ invariant under permutations,

$$E\phi(Z_1^2, \ldots, Z_N^2) \leq E\phi(\tilde{Z}_1^2, \ldots, \tilde{Z}_N^2) .$$

(Marshall-Proschan)

$\mathbf{Y} := (Y_1, \ldots, Y_N)^*$, $\mathbf{X} = T\mathbf{Y}$, want: $E(\max_i X_i^2)$ is maximal when $T = I_N$.

Sketch of Proof: It is enough to prove $E(\max_i X_i^2) \leq E(\max_i Y_i^2)$.

Let $\tilde{X}$ be a vector of independent Gaussian with $EX_i^2 = E\tilde{X}_i^2$. By Sidak's inequality,

$$E(\max_i X_i^2) \leq E(\max_i \tilde{X}_i^2)\,.$$

The function $\phi(\mathbf{X}) = \max X_i^2$ is convex and invariant under permutations. Further, the variances of the Y_i dominate those of X_i (and hence those of $\tilde{X}_i$)

Because $\sum_j T_{1j}^2 = 1$, $EX_1^2 = \sum_j T_{1,j}^2 EY_j^2 \le EY_1^2$,

$$\sum_{i=1}^{k+1} EX_i^2 = \sum_j \sum_{i=1}^{k+1} T_{i,j}^2 EY_j^2$$

$$= \sum_{j=1}^{k} \sum_{i=1}^{k+1} T_{i,j}^2 EY_j^2 + \sum_{j=k+1}^{N} \sum_{i=1}^{k+1} T_{i,j}^2 EY_j^2$$

$$\le \sum_{j=1}^{k} \sum_{i=1}^{k+1} T_{i,j}^2 EY_j^2 + EY_{k+1}^2 \sum_{j=k+1}^{N} \sum_{i=1}^{k+1} T_{i,j}^2$$

$$= \sum_{j=1}^{k} EY_j^2 \sum_{i=1}^{k+1} T_{i,j}^2 + EY_{k+1}^2 [k + 1 - \sum_{j=1}^{k} \sum_{i=1}^{k+1} T_{i,j}^2]$$

$$= \sum_{j=1}^{k+1} EY_j^2 + \sum_{j=1}^{k} [EY_{k+1}^2 - EY_j^2][1 - \sum_{i=1}^{k+1} T_{i,j}^2]$$

$$\le \sum_{j=1}^{k+1} EY_j^2 .$$

Since also $\sum_i EY_i^2 = \sum_i EX_i^2 = \sum_i E\tilde{X}_i^2$, by the Marshall-Proschan inequality,

$$E(\max_i \tilde{X}_i^2) \leq E(\max_i Y_i^2)\,,$$

which completes the proof.

Remarks: 1) The Schur convexity part works also for the function

$$\phi_k(\mathbf{x}) = \max_{i_1 \neq i_2 \neq \ldots \neq i_k} \sum_{j=1}^{k} x_{i_j}^2 \, .$$

What is missing in order to prove the conjecture for general k is the analog of Sidak's inequality: is it true that

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

Note that this does NOT follow from the Gaussian correlation inequality.

$$\phi_k(\mathbf{x}) = \max_{i_1 \neq i_2 \neq \ldots \neq i_k} \sum_{j=1}^{k} x_{i_j}^2 .$$

2) Gordon, Litvak, Schutt and Werner (AP 2002) proved that

$$E(\phi_k(\mathbf{X})) \leq \frac{4e}{e-1} E(\phi_k(\tilde{\mathbf{X}})) .$$

Unfortunately, this only proves that

$$\left(E\|\mathbf{X}\|_2^2 - \mathcal{E}(N,k,T)\right) \leq \frac{4e}{e-1}\left(E(\|\mathbf{X}\|_2^2) - \mathcal{E}(N,k,I_N)\right)$$

This does not give control on the relative error.

3) If $k = N-1$, one needs (after appealing to the Schur convexity part of the argument) consider the expectation

$$E \max_{i_1 \neq i_2 \neq \ldots \neq i_{N-1}} \sum_{j=1}^{k} X_{i_j}^2 = E\|X\|_2^2 - E \min_i X_i^2 \,.$$

Another inequality of Gordon, Litvak, Schutt and Werner (CRAS 2005) shows that

$$E \min_i X_i^2 \geq \frac{1}{6} E \min_i \tilde{X}_i^2 \,.$$

It then follows that

$$\mathcal{E}(N, k, I_N) \leq 6\mathcal{E}(N, k, T) \,.$$

That is, lost at most a factor 6 in the error.

4) Depending on the rate of decay of the eigenvalues of $\mathbf{R}$, various asymptotics can be studied. We have not systematically studied that.