Mathematics > Probability

Covariance Estimation for Distributions with $2+\varepsilon$ Moments

Nikhil Srivastava, Roman Vershynin

(Submitted on 14 Jun 2011 (v1), last revised 30 Mar 2012 (this version, v3))

We study the minimal sample size N=N(n) that suffices to estimate the covariance matrix of an n-dimensional distribution by the sample covariance matrix in the operator norm, and with an arbitrary fixed accuracy. We establish the optimal bound N = O(n) for every distribution whose k-dimensional marginals have uniformly bounded 2+\epsilon moments outside the sphere of radius O(\sqrt{k}). In the specific case of log-concave distributions, this result provides an alternative approach to the Kannan-Lovasz-Simonovits problem, which was recently solved by Adamczak, Litvak, Pajor and Tomczak-Jaegermann. Moreover, a lower estimate on the covariance matrix holds under a weaker assumption -- uniformly bounded 2+\epsilon moments of one-dimensional marginals. Our argument proceeds by randomizing the spectral sparsification technique of Batson, Spielman and Srivastava. The spectral edges of the sample covariance matrix are controlled via the Stieltjes transform evaluated at carefully chosen random points.

Comments:	26 pages
Subjects:	**Probability (math.PR)**; Statistics Theory (math.ST)
MSC classes:	60B20, 15B52 (Primary), 62H12 (Secondary)
Cite as:	**arXiv:1106.2775 [math.PR]**
	(or **arXiv:1106.2775v3 [math.PR]** for this version)

Submission history

From: Nikhil Srivastava [view email]
[v1] Tue, 14 Jun 2011 18:14:35 GMT (23kb)
[v2] Thu, 1 Mar 2012 08:34:51 GMT (24kb)
[v3] Fri, 30 Mar 2012 19:41:24 GMT (24kb)

Which authors of this paper are endorsers?