Invariance principle on the slice. (English summary)


The Berry-Esseen theorem gives bounds on the speed of convergence of a sum $\sum_i X_i$ to the corresponding Gaussian distribution. Convergence occurs as long as none of the summands $X_i$ are too "prominent". E. Mossel, R. O'Donnell and K. Oleszkiewicz [Ann. of Math. (2) 171 (2010), no. 1, 295–341; MR2630040] proved the non-linear invariance principle, an analog of the Berry-Esseen theorem for low-degree polynomials:

Given a low-degree polynomial $f$ on $n$ variables in which none of the variables is too prominent (technically, $f$ has low influences), the invariance principle states that the distribution of $f(X_1, \ldots, X_n)$ and $f(Y_1, \ldots, Y_n)$ is similar as long as (1) each of the vectors $(X_1, \ldots, X_n)$ and $(Y_1, \ldots, Y_n)$ consists of independent coordinates, (2) the distributions of $X_i$, $Y_i$ have matching first and second moments, and (3) the variables $X_i$, $Y_i$ are hypercontractive.

In particular, the result establishes that if $f(x_1, \ldots, x_n)$ is a multilinear low-degree polynomial with low influences, then the distribution of $f(X_1, \ldots, X_n)$ is close (in various senses) to the distribution of $f(G_1, \ldots, G_n)$, where $B_i$ are independent Bernoulli random variables and $G_i$ are independent standard Gaussians.

The invariance principle comes up in the context of proving the Majority is the Stablest conjecture (which shows that the Goemans-Williamson algorithm for MAX-CUT is optimal under the Unique Games Conjecture). In general, the principle allows the analysis of a function on the Boolean cube (corresponding to the $X_i$) by analyzing its counterpart in Gaussian space (corresponding to the $Y_i$), in which geometric methods can be applied.

Here the authors prove the invariance principle for the distribution over $X_1, \ldots, X_n$ which is uniform over the slice

$$\binom{[n]}{k} = \{ (x_1, \ldots, x_n) \in \{0, 1\}^n : x_1 + \cdots + x_n = k \}.$$ 

If $f$ is a low-degree function on $\binom{[n]}{k}$ having low influences, then the distributions of $f(X_1, \ldots, X_n)$ and $f(Y_1, \ldots, Y_n)$ are close, where $X_1, \ldots, X_n$ is the uniform distribution on $\binom{[n]}{k}$, and $Y_1, \ldots, Y_n$ are either independent Bernoulli variables with expectation $k/n$, or independent Gaussians with the same mean and variance.

This setting arises naturally in hardness of approximation and in extremal combinatorics (the Erdős-Ko-Rado theorem and its many extensions).

The Majority is Stablest conjecture and Bourgain’s tail bound can be generalized to this setting. Also, using Bourgain’s tail bound, the authors prove an analog of the Kindler-Safra theorem, which states that if a Boolean function is close to a function of constant degree, then it is close to a junta. Finally, as a corollary of the new version of the Kindler-Safra theorem, the authors prove a stability version of the $t$-intersecting Erdős-Ko-Rado theorem. E. Friedgut [Combinatorica 28 (2008), no. 5, 503–528; MR2501247] showed that for all $\lambda, \zeta > 0$ there exists $\epsilon = \epsilon(\lambda, \zeta) > 0$ such that for all $k, n$ satisfying

$$\lambda < \frac{k}{n} < \frac{1}{t+1} - \zeta,$$
a \( t \)-intersecting family in \( \binom{n}{k} \) of almost maximal size \((1 - \epsilon)\binom{n-t}{k-t}\) is close to an optimal family (a \( t \)-star). The authors extend this result to the regime \( k/n \approx 1/(t + 1) \). (When \( k/n > 1/(t + 1) \), \( t \)-stars are no longer optimal.)

The proof of the main result combines algebraic, geometric, and analytic ideas. A coupling argument, which crucially relies on properties of harmonic functions, shows that the distribution of a low-degree, low-influence harmonic function is approximately invariant moving from the original slice to nearby slices. Taken together, these slices form a thin layer around the original slice, on which the function has roughly the same distribution as on the original slice. The classical invariance principle implies that the distribution of the given function on the layer is close to its distribution on the Gaussian counterpart of the layer, which turns out to be identical to its distribution on all of Gaussian space, completing the proof.

Y. Filmus and Mossel [in 31st Conference on Computational Complexity, Art. No. 16, 13 pp., LIPIcs. Leibniz Int. Proc. Inform., 50, Schloss Dagstuhl. Leibniz-Zent. Inform., Wadern, 2016; MR3540817] have shown that a special case of the main result in the paper can be obtained under more relaxed assumptions on the polynomials of interest. The main result in that paper (which subsumes the main result of this article) uses completely different proof techniques, in particular without recourse to Gaussian spaces.

Michele Zito

References


*Note: This list reflects references listed in the original paper as accurately as possible with no attempt to correct errors.*

© Copyright American Mathematical Society 2021