A SIMPLE DECOUPLING INEQUALITY IN PROBABILITY THEORY

ROMAN VERSHYNIN

Abstract. We present a basic decoupling inequality in probability theory. Both the result and its proof are well known, but this short proof is not easy to locate in the literature.

Decoupling is a technique of replacing quadratic forms of random variables by bilinear forms. The monograph [2] offers a systematic study of decoupling and its applications. In this note, we state and prove a simple and useful decoupling inequality. In a more general form, for multilinear forms, this inequality can be found in [2, Theorem 3.1.1]; see also [1, Theorem 8.11] for a similar inequality for quadratic forms.

Theorem 1. *Let A be an $n \times n$ matrix with zero diagonal. Let $X = (X_1, \ldots, X_n)$ be a random vector with independent mean zero coefficients. Then, for every convex function F, one has*

$$\mathbb{E}\, F(\langle AX, X \rangle) \le \mathbb{E}\, F(4 \langle AX, X' \rangle)$$

where X' is an independent copy of X.

The consequence of the theorem can be equivalently stated as

$$\mathbb{E}\, F\Big(\sum_{i,j=1}^n a_{ij} X_i X_j \Big) \le \mathbb{E}\, F\Big(4 \sum_{i,j=1}^n a_{ij} X_i X'_j \Big)$$

where $X' = (X'_1, \ldots, X'_n)$.

Proof. Let $A = (a_{ij})_{i,j=1}^n$, and let $\delta_1, \ldots, \delta_n$ be independent Bernoulli random variables with $\mathbb{P}\{\delta_i = 0\} = \mathbb{P}\{\delta_i = 1\} = 1/2$. We shall denote the conditional expectation with respect to these δ_i by $\mathbb{E}_\delta$, and similarly for the conditional expectations with respect to X and X'. We express

$$\langle AX, X \rangle = \sum_{i,j \in [n]} a_{ij} X_i X_j = 4\, \mathbb{E}_\delta \sum_{i,j \in [n]} \delta_i (1 - \delta_j) a_{ij} X_i X_j.$$

By Jensen's and Fubini inequalities,

$$\mathbb{E}\, F(\langle AX, X \rangle) \le \mathbb{E}_\delta\, \mathbb{E}_X\, F\Big(4 \sum_{i,j \in [n]} \delta_i (1 - \delta_j) a_{ij} X_i X_j \Big)$$

Date: May 10, 2011.

Let us fix a realization of $\delta_1, \dots, \delta_n$ and consider the subset $I = \{i \in [n] : \; \delta_i = 1\}$. We have

$$4 \sum_{i,j \in [n]} \delta_i (1-\delta_j) a_{ij} X_i X_j = 4 \sum_{(i,j) \in I \times I^c} a_{ij} X_i X_j.$$

Since X_i, $i \in I$ are independent from X_j, $j \in I^c$, the distribution of this sum will not change if we replace X_j by X'_j, the coordinates of X'. So

$$\mathbb{E}\, F(\langle AX, X \rangle) \le \mathbb{E}_\delta \, \mathbb{E}_{X,X'} \, F\Big(4 \sum_{(i,j) \in I \times I^c} a_{ij} X_i X'_j \Big)$$

Finally, we use a simple consequence of Jensen's inequality: if Y and Z are independent random variables and $\mathbb{E}\, Z = 0$ then $\mathbb{E}\, F(Y) = \mathbb{E}\, F(Y + \mathbb{E}\, Z) \le \mathbb{E}(Y+Z)$. Using this fact for $Y = 4 \sum_{(i,j) \in I \times I^c} a_{ij} X_i X'_j$ and $Z = 4 \sum_{(i,j) \notin I \times I^c} a_{ij} X_i X'_j$, we obtain

$$\mathbb{E}_{X,X'} \, F(Y) = \mathbb{E}_{X,X'} \, F(Y+Z) = \mathbb{E}\, F(4 \langle AX, X \rangle).$$

Taking the expectation with respect to (δ_i), we complete the proof. $\square$

References

[1] S. Foucart and H. Rauhut, *A mathematical introduction to compressive sensing.* Springer, 2013

[2] V. de la Pena, E. Gine, *Decoupling. From dependence to independence. Randomly stopped processes. U-statistics and processes. Martingales and beyond.* Probability and its Applications (New York). Springer-Verlag, New York, 1999.

Department of Mathematics, University of Michigan, 530 Church St., Ann Arbor, MI 48109, U.S.A.

E-mail address: `romanv@umich.edu`