Computer Science 294 Lecture 21 Notes

Daniel Raban

April 4, 2023

1 The Invariance Principle

1.1 Comparing random bits and Gaussians

Recall the Berry-Essen theorem, a quantitative version of the Central Limit Theorem:

Theorem 1.1 (Berry-Essen)**.** *Let $X_1, \dots, X_n$ be independent random variables with $\mathbb{E}[X_i] = \mathbb{E}[X_i^3] = 0$ and $\mathbb{E}[X_i^2] = \sigma_i^2$, where $\sum_{i=1}^n \sigma_i^2 = 1$. Let $S = X_1 + \cdots + X_n$ and $Z \sim N(0,1)$. Then the CDF of S is $O(\beta)$-close to the CDF of Z, where $\beta = \sum_{i=1}^m \mathbb{E}[X_i^4]$.*

The invariance principle relates the analysis of Boolean functions to the analysis of functions over Gaussians. If we think of $X_1, \dots, X_n$ as independent random bits, then the Berry-Essen theorem is a basic example of this principle.

Example 1.1. Let $x_1, \dots, x_n \sim \{\pm 1\}^n$, and let

$$X_1 = a_1 x_1, \qquad X_2 = a_2 x_2, \qquad \dots \qquad X_n = a_n x_n.$$

If we take $S = \sum_{i=1}^n a_i x_i$, then $0 = \mathbb{E}[X_i] = \mathbb{E}[X_i^3]$ with $\mathbb{E}[X_i^2] = a_i^2 \, \mathbb{E}[x_i^2] = a_i^2$. Assume that $\sum_{i=1}^n a_i^2 = 1$. Then

$$\beta = \sum_{i=1}^n \mathbb{E}[X_i^4] = \sum_{i=1}^n a_i^4 \leq \max_i a_i^2 \sum_{i=1}^n a_i^2 = \max_i a_i^2.$$

If $a_1 = a_2 = \cdots = a_n = \frac{1}{\sqrt{n}}$, then $\beta = 1/n$. Then

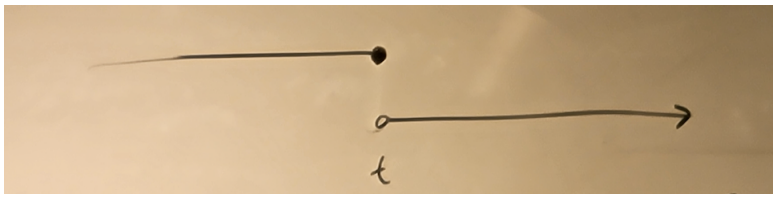
$$\frac{x_1 + \cdots + x_n}{\sqrt{n}} \approx_{1/n} Z.$$

Rather than comparing the CDFs, it will be more useful for us to work with the following notion of "closeness" of distributions: For any "nice" test function ψ,

$$\mathbb{E}[\psi(S)] \approx \mathbb{E}[\psi(X)].$$

Example 1.2. One class of test functions would be

$$\psi_t(x) = \mathbb{1}_{\{x \le t\}}.$$



This corresponds to CDF closeness:

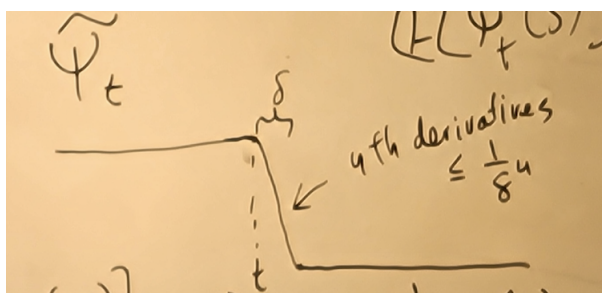
$$\mathbb{P}(S \le t) = \mathbb{E}[\psi_t(S)] \approx \mathbb{E}[\psi_t(Z)] = \mathbb{P}(Z \le t).$$

Theorem 1.2. *Let $X_1, \dots, X_n$ be independent random variables with $\mathbb{E}[X_i] = \mathbb{E}[X_i^3] = 0$ and $\mathbb{E}[X_i^2] = \sigma_i^2$, where $\sum_{i=1}^n \sigma_i^2 = 1$. Let $S = X_1 + \cdots + X_n$ and $Z \sim N(0,1)$, and let $\psi \in C^4$ with bounded 4th derivative. Then*

$$|\,\mathbb{E}[\psi(S)] - \mathbb{E}[\psi(Z)]| \le \|\psi^{(4)}\|_\infty \cdot O(\beta),$$

where $\beta = \sum_{i=1}^m \mathbb{E}[X_i^4]$.

Remark 1.1. To deal with non-smooth test functions like the example above, we can approximate ψ_t by $\widetilde{\psi}_t$, where $\widetilde{\psi}_t$ is a smoothed version of ψ_t with 4th derivative $\le 1/\delta^4$:



This gives

$$|\,\mathbb{E}[\psi_t(S)] - \mathbb{E}[\psi_t(Z)]| \le O(\delta) + O(\beta) \cdot \frac{1}{\delta^4} \le O(\delta),$$

and we can then pick $\delta = \beta^{1/5}$.

Why do we need to care about the size of the derivatives? To differentiate $\frac{x_1 + \cdots + x_n}{\sqrt{n}}$ from $Z \sim N(0,1)$, we could try to use rapidly oscillating functions such as $\psi_n(x) = \cos(\sqrt{n} \cdot 2\pi x)$. Then we would have $\mathbb{E}[\psi(S)] = 1$ and $\mathbb{E}[\psi_n(Z)] \xrightarrow{n \to \infty} 0$.

Proof. To compare the sum $X_1 + \cdots + X_n$ of random variables to $Z \sim N(0,1)$, we can write $Z = Z_1 + \cdots + Z_n$, where the Z_i are independent Gaussians with $Z_i \sim N(0, \sigma_i^2)$. Note that $\mathbb{E}[Z_i] = \mathbb{E}[X_i] = 0$, $\mathbb{E}[Z_i^2] = \sigma_i^2 = \mathbb{E}[X_i^2]$, and $\mathbb{E}[Z_i^3] = 0$ for all i.

How can we replace $X_1 + \cdots + X_n$ with $Z_1 + \cdots + Z_n$? One way is to replace one X_i with a Z_i at a time.[1] Define the hybrid random variables

$$H_i = Z_1 + \cdots + Z_i + X_{i+1} + \cdots X_n.$$

Thus,

$$H_0 = X_1 + \cdots X_n = S, \qquad H_n = Z_1 + \cdots + Z_n = Z.$$

Now we can say

$$\begin{aligned} |\,\mathbb{E}[\psi(S)] - \mathbb{E}[\psi(Z)] &= |\,\mathbb{E}[\psi(H_0)] - \mathbb{E}[\psi(H_n)] \\ &= \left| \sum_{i=1}^{n} \mathbb{E}[\psi(H_{i-1})] - \mathbb{E}[\psi(H_i)] \right| \end{aligned}$$

Using the triangle inequality,

$$\leq \sum_{i=1}^{n} |\,\mathbb{E}[\psi(H_{i-1})] - \mathbb{E}[\psi(H_i)]|$$

H_{i-1} and H_i only differ in the i-th summand. In particular, we can write $H_{i-1} = U + Z_i$ and $H_i = U + X_i$, where $U = Z_1 + \cdots Z_{i-1} + X_{i+1} \cdots + X_n$ is independent of Z_i and X_i. So it suffices to show that $|\,\mathbb{E}[\psi(U + X_i)] - \mathbb{E}[\psi(U + Z_i)]|$ is small. The idea, then, is to Taylor expand ψ around U.

Fix the value $U = u$, and write

$$\psi(u + \delta) = \psi(u) + \psi'(u) \cdot \delta + \psi''(u) \cdot \frac{\delta^2}{2!} + \psi'''(u) \cdot \frac{\delta^3}{3!} + \psi^{(4)}(u^*) \frac{\delta^4}{4!},$$

where u^* is between u and $u + \delta$. Now we can write

$$\psi(U + X_i) = \psi(U) + \psi'(U) \cdot X_i + \psi''(U) \cdot \frac{X_i^2}{2!} + \psi'''(U) \cdot \frac{X_i^3}{3!} + \psi^{(4)}(u^*) \frac{X_i^4}{4!},$$

$$\psi(U + Z_i) = \psi(U) + \psi'(U) \cdot Z_i + \psi''(U) \cdot \frac{Z_i^2}{2!} + \psi'''(U) \cdot \frac{Z_i^3}{3!} + \psi^{(4)}(u^{**}) \frac{Z_i^4}{4!},$$

where U^* is between U and $U + X_i$ and U^{**} is between U and $U + Z_i$. Now, using the linearity of expectation,

$$|\,\mathbb{E}[\psi(U + X_i) - \psi(U + Z_i)]| = \left| \mathbb{E}\left[\psi^{(4)}(U^*) \frac{X_i^4}{4!}\right] - \mathbb{E}\left[\psi^{(4)}(U^{**}) \frac{Z_i^4}{4!}\right] \right|$$

[1] This is known as the **replacement method**. In computer science, this is called the **hybrid argument**.

$$\leq \left| \mathbb{E}\left[\psi^{(4)}(U^*)\frac{X_i^4}{4!}\right]\right| + \left| \mathbb{E}\left[\psi^{(4)}(U^{**})\frac{Z_i^4}{4!}\right]\right|$$
$$\leq \|\psi^{(4)}\|_\infty \frac{1}{4!}\,\mathbb{E}[X_i^4] + \|\psi^{(4)}\|_\infty \frac{1}{4!}\,\mathbb{E}[Z_i^4]$$

Using a property of Gaussian moments,

$$= \|\psi^{(4)}\|_\infty \frac{1}{4!}\,\mathbb{E}[X_i^4] + \|\psi^{(4)}\|_\infty \frac{3}{4!}(\mathbb{E}[Z_i^2])^2$$

Since $\mathbb{E}[Z_i^2] = \mathbb{E}[X_i^2]$.

$$= \|\psi^{(4)}\|_\infty \frac{1}{4!}\,\mathbb{E}[X_i^4] + \|\psi^{(4)}\|_\infty \frac{3}{4!}\|X_i\|_2^4$$

Since $\|X_i\|_2 \leq \|X_i\|_4$,

$$\leq \|\psi^{(4)}\|_\infty \frac{4}{4!}\,\mathbb{E}[X_i^4].$$

In total, we have

$$|\,\mathbb{E}[\psi(S)] - \mathbb{E}[\psi(Z)]| \leq \sum_{i=1}^n |\,\mathbb{E}[\psi(H_{i-1})] - \mathbb{E}[\psi(H_i)]|$$
$$\leq \sum_{i=1}^n \|\psi^{(4)}\|_\infty \cdot \frac{4}{4!}\,\mathbb{E}[X_i^4],$$

which completes the proof. □

Really, our proof has proven the following:

Theorem 1.3. *Let $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ be independent random variables with $\mathbb{E}[X_i^j] = \mathbb{E}[Y_i^j]$ for $1 \leq i \leq n$ and $1 \leq j \leq 3$. Let $S_X := \sum_{i=1}^n X_i$ and $S_Y := \sum_{i=1}^n Y_i$. Then*

$$|\,\mathbb{E}[\psi(S_X)] - \mathbb{E}[\psi(S_Y)]| \leq \frac{1}{24}\|\psi^{(4)}\|_\infty \cdot \beta_{X,Y},$$

where $\beta_{x,Y} = \sum_{i=1}^n \mathbb{E}[X_i^4] + \mathbb{E}[Y_i^4]$.

1.2 The invariance principle

Here is an extension of our previous result.

Theorem 1.4. *If $X_1, \ldots, X_n$ and $Y_1, \ldots, Y_n$ are independent $\mathbb{R}^d$-valued random variables with matching first and second moments and ψ as bounded 3rd derivatives, then*

$$\mathbb{E}\left[\psi\left(\sum_{i=1}^n X_i\right)\right] \approx \mathbb{E}\left[\psi\left(\sum_{i=1}^n Y_i\right)\right].$$

Here is a second extension.

Theorem 1.5 (Invariance principle)**.** *Let $f : \mathbb{R}^n \to \mathbb{R}$ be a multilinear polynomial of degree d, i.e.*

$$f(x) = \sum_{S \subseteq [n]} \widehat{f}(S) \prod_{i \in S} x_i.$$

Let $X_1, \ldots, X_n \sim \{\pm 1\}$ be independent random bits, and let $Y_1, \ldots, Y_n \sim N(0,1)$ be independent standard Gaussians. Then

$$|\,\mathbb{E}[\psi(f(X_1, \ldots, X_n))] - \mathbb{E}[\psi(f(Y_1, \ldots, Y_n))]| \leq \frac{\|\psi^{(4)}\|_\infty}{24} \cdot 9^{d-1} \cdot \sum_{i=1}^{n} \mathrm{Inf}_i^2(f)(\mathbb{E}[X_i^4] + \mathbb{E}[Y_i^4]),$$

where $\mathrm{Inf}_i(f) = \sum_{S \ni i} \widehat{f}(S)^2$.

Remark 1.2. Berry-Essen is the special case of this theorem with $f(x) = \sum_{i=1}^n a_i x_i$ with $\sum_i a_i^2 = 1$ and $\mathrm{Inf}_i(f) = a_i^2$.

Before we prove the theorem, let's note that our Fourier analysis holds for polynomials $f : \mathbb{R}^n \to \mathbb{R}$, not just $f : \{\pm 1\}^n \to \mathbb{R}$.

Proposition 1.1 (Plancherel's identity)**.** *Let $f, g : \mathbb{R}^n \to \mathbb{R}$ be two multilinear polynomials:*

$$f(x) = \sum_{S \subseteq [n]} \widehat{f}(S) \prod_{i \in S} x_i, \qquad g(x) = \sum_{S \subseteq [n]} \widehat{g}(S) \prod_{i \in S} x_i.$$

Let $Z_1, \ldots, Z_n$ be any independent random variable with mean 0 and variance 1. Then

$$\mathbb{E}[f(Z)g(Z)] = \sum_{S \subseteq [n]} \widehat{f}(S)\widehat{g}(S).$$

Proof.

$$\begin{aligned}
\mathbb{E}_Z[f(Z)g(Z)] &= \mathbb{E}\left[\left(\sum_{S \subseteq [n]} \widehat{f}(S) \prod_{i \in S} Z_i\right)\left(\sum_{T \subseteq [n]} \widehat{g}(T) \prod_{i \in T} Z_i\right)\right] \\
&= \sum_{S,T} \widehat{f}(S)\widehat{g}(T)\, \mathbb{E}\left[\prod_{i \in S} Z_i \cdot \prod_{i \in T} Z_i\right] \\
&= \sum_{S,T} \widehat{f}(S)\widehat{g}(T)\, \mathbb{E}\left[\prod_{i \in S \cap T} Z_i^2 \cdot \prod_{i \in S \triangle T} Z_i\right]
\end{aligned}$$

Since the Z_i are indepedent,

$$\begin{aligned}
&= \sum_{S,T} \widehat{f}(S)\widehat{g}(T)\, \mathbb{E}\left[\prod_{i \in S \cap T} Z_i^2\right] \cdot \mathbb{E}\left[\prod_{i \in S \triangle T} Z_i\right] \\
&= \sum_{S} \widehat{f}(S)\widehat{g}(S).
\end{aligned}$$

□

Corollary 1.1 (Parseval's identity)**.**

$$\mathbb{E}[f(Z)^2] = \sum_S \widehat{f}(S)^2.$$

Proposition 1.2 (Inversion formula)**.**

$$\mathbb{E}[f(Z) \cdot 1] = \widehat{f}(\varnothing), \qquad \mathbb{E}\left[f(Z) \prod_{i \in S} Z_i\right] = \widehat{f}(S).$$

Define the derivative $D_i f(x) = \sum_{S \in i} \widehat{f}(S) \prod_{j \in S \setminus \{i\}} x_j$. Then

$$\mathbb{E}[D_i f(Z)^2] = \sum_{S \ni i} \widehat{f}(S)^2 =: \mathrm{Inf}_i(f).$$

Bonami's lemma still holds, as well, as long as the Z_i are 9-reasonable.

Proof sketch of invariance principle. We want to show

$$\mathbb{E}_{X_1,\ldots,X_n \sim \{\pm 1\}}[\psi(f(X_1, \ldots, X_n))] \approx \mathbb{E}_{Z_1,\ldots,Z_n \sim N(0,1)}[\psi(f(Z_1, \ldots, Z_n))],$$

so define the hybrids

$$H_i = f(Z_1, \ldots, Z_i, X_{i+1}, \ldots, X_n).$$

As before, it suffices to show that for all i, $\mathbb{E}[\psi(H_{i-1})] \approx \mathbb{E}[\psi(H_i)]$. We can write

$$f(x) = x_i D_i f(x) + E_i f(x),$$

where $D_i f(X)$ and $E_i f(X)$ are independent of X_i. Since H_i and H_{i-1} only differ in the i-th coordinate, we have

$$f(H_i) = Z_i D_i f(Z_1, \ldots, Z_{i-1} X_{i+1}, \ldots, X_n) + E_i f(Z_1, \ldots, Z_{i-1}, X_{i+1}, \ldots, X_n),$$

$$f(H_{i-1}) = X_i D_i f(Z_1, \ldots, Z_{i-1} X_{i+1}, \ldots, X_n) + E_i f(Z_1, \ldots, Z_{i-1}, X_{i+1}, \ldots, X_n),$$

Now write

$$f(H_i) = Z_i \cdot \Delta + U, \qquad f(H_{i-1}) = X_i \cdot \Delta + U.$$

We will finish the proof sketch next time. $\square$