The MMSE Conjecture

Michel Ledoux

University of Toulouse, France

Abstract

This note presents the MMSE conjecture to attract possible interest.

Let f be a probability density on the real line. Provided there are well-defined, consider the entropy of f,

$$\mathrm{H}(f) \,=\, \int_{\mathbb{R}} f \log f \, dx$$

(up to the sign convention) and its Fisher information

$$\mathrm{I}(f) \,=\, \int_{\mathbb{R}} \frac{f'^2}{f} \, dx.$$

Denote by $p_t(x)$ the Gaussian kernel

$$p_t(x) \,=\, \frac{1}{\sqrt{2\pi t}} \, \mathrm{e}^{-x^2/2t}, \quad t > 0, \ x \in \mathbb{R},$$

and let $f_t = f * p_t$, $t > 0$.

The MMSE conjecture basically asks whether the knowledge of $\mathrm{H}(f_t)$, $t > 0$, or $\mathrm{I}(f_t)$, $t > 0$, is enough to determine f.

To precisely describe the conjecture (and the terminology), let us slightly extend the picture.

Let μ be a probability measure on the Borel sets of $\mathbb{R}$. For every $t > 0$, let μ_t be the convolution of μ with the Gaussian kernel p_t. That is, μ_t has the strictly positive C^∞ density f_t with respect to the Lebesgue measure which is given as the convolution

$$f_t(x) \,=\, \int_{\mathbb{R}} p_t(x-y) d\mu(y), \quad x \in \mathbb{R}.$$

From a probabilistic perspective, let X be a random variable on a probability space $(\Omega, \mathcal{A}, \mathbb{P})$ with law μ. Then μ_t is the law of $X_t = X + \sqrt{t}\, N$ where N is a standard normal variable independent of X.

Now, $\mathrm{I}(f_t)$ is well-defined for every $t > 0$ (see below) and we may ask whether its knowledge determines μ. The classical de Brujin formula expresses that

$$\frac{d}{dt}\,\mathrm{H}(f_t) \,=\, -\frac{1}{2}\,\mathrm{I}(f_t)$$

so that the question may be formulated equivalently for the entropy.

In the sequel, write $\mathrm{I}(\mu_t) = \mathrm{I}(X_t) = \mathrm{I}(f_t)$, $t > 0$.

It is immediate to see that the quantities $\mathrm{I}(\mu_t)$, $t > 0$, are invariant by translation and symmetry. That is, if ν is the distribution defined by

$$\int_{\mathbb{R}} \varphi\, d\nu \,=\, \int_{\mathbb{R}} \varphi(\pm x + a) d\mu(x)$$

for every bounded measurable $\varphi : \mathbb{R} \to \mathbb{R}$, where $a \in \mathbb{R}$, then $\mathrm{I}(\nu_t) = \mathrm{I}(\mu_t)$ for every $t > 0$. Equivalently, $\mathrm{I}(X_t)$, $t > 0$, is invariant by the change of X into $\pm X + a$ for any $a \in \mathbb{R}$.

Conjecture. *Let μ be a probability distribution on $\mathbb{R}$. Then the knowledge of* $\mathrm{I}(\mu_t)$*, $t > 0$, characterizes μ up to translation and symmetry.*[1]

In probabilistic terms, does $\mathrm{I}(X_t)$, $t > 0$, characterizes the law of X up to the change of X into $X + a$ or $-X$?

This conjecture came up in information theory in the work [2] by D. Guo, Y. Wu, S. Shamai and S. Verdú on the estimation of a random variable from its observation perturbated by a Gaussian noise. More precisely, in the preceding notation, consider

$$\mathrm{MMSE}(t) \,=\, \mathbb{E}\Big(\big[X - \mathbb{E}\big(X \,|\, X_t\big)\big]^2\Big), \quad t \geq 0,$$

which is known as the Minimum Mean-Square Error (MMSE). Note that although $\mathbb{E}(X \,|\, X_t)$ may not be well-defined if X is not integrable, it makes sense to consider the integrable random variable $X_t - \mathbb{E}(X \,|\, X_t)$ which is identified to $\sqrt{t}\,\mathbb{E}(N \,|\, X_t)$ since

$$X_t \,=\, \mathbb{E}\big(X_t \,|\, X_t\big) \,=\, \mathbb{E}\big(X \,|\, X_t\big) + \sqrt{t}\,\mathbb{E}\big(N \,|\, X_t\big).$$

In particular, $X - \mathbb{E}(X \,|\, X_t)$ makes also sense and has moments of all orders.

The MMSE connects to the Fisher information $\mathrm{I}(X_t)$, $t > 0$, along the heat flow via the identity

$$t^2\,\mathrm{I}(X_t) \,=\, t - \mathrm{MMSE}(t), \quad t > 0.$$

(cf. [1, 3]). (Note that the invariances by translation and symmetry are immediate on this representation.)

Some partial results on the conjecture are described in [1, 2, 3]. The multi-dimensional case may also be discussed.

[1]When I mentioned this conjecture to a close colleague, I got the following answer: to characterize a probability distribution, there is a convenient tool, the Fourier transform.

References

[1] D. Guo, S. Shamai, S. Verdú. Mutual information and minimum mean-square error in Gaussian channels. *IEEE Trans. Inform. Theory* 51 (2005), 1261–1282.

[2] D. Guo, Y. Wu, S. Shamai, S. Verdú. Estimation in Gaussian noise: properties of the minimum mean-square error. *IEEE Trans. Inform. Theory* 57 (2011), 2371–2385.

[3] M. Ledoux. Heat Flow Derivatives and Minimum Mean-Square Error in Gaussian Noise. *IEEE Trans. Inform. Theory*, to appear (2016).