arXiv:2303.09023v1 [math.ST] 16 Mar 2023

最不利噪声的存在性

Dongzhou Huang*

2023年3月17日

摘要

设观测到一个感兴趣的随机变量 X。本文研究“最不利噪声” $\hat{}\ Y_\epsilon$，它使预测误差 $E[X - E[X|X+Y]]^2$(或使 $E[X|X+Y]$ 的方差最小) 在满足以下条件的 Y 类中取最大值： Y与X 独立，且方差为$Y \le \epsilon^2$。Ernst、Kagan 和 Rogers（[3]）最早研究了这个问题。$\hat{}$ 在本文中，我们证明最不利噪声 Y_ϵ 必定存在，且其方差必定为 ϵ^2。存在性证明依赖于我们针对条件期望的方差所建立的一个收敛结果。此外，我们证明函数 $\inf_{\text{var}Y \le \epsilon^2}$ 方差 $E[X|X+Y]$ 关于 ϵ严格递减且右连续。

1 引言

2022 年，Ernst、Kagan 和 Rogers（[3]）研究了观测到平方可积随机变量 X 时“最不利噪声”的问题。作者考虑了 Y_ϵ，这是一个与 X 独立的平方可积随机变量，它使预测误差最大

$$E\left[X - E[X|X+Y]\right]^2 = \text{var}X - \text{var}\, E[X|X+Y],$$

（或者等价地，使 $E[X|X+Y]$的方差最小），且属于方差为 var$Y \le \epsilon^2$ 的 Y 类。作者随后刻画了最不利噪声，并证明如果 Y 的分布满足某一特定刻画的条件，那么 Y 就应是最不利噪声。

本文从 Ernst 等人对最不利噪声的刻画中退后一步，提出了以下问题：“最不利噪声是否存在？”换言之，给定 X的分布，是否存在一个 Y的分布，使 $E\left[X - E[X|X+Y]\right]^2$ 达到最大值？尽管 Ernst 等人证明了在以下条件下最不利噪声存在：(i) X 的分布已知；(ii) 给定该分布

*统计系，科罗拉多州立大学，美利坚合众国。电子邮件：dongzhou.huang@colostate.edu

Y 满足 [3], 中的三个刻画条件，而这些条件较为复杂，更重要的是，对于大多数 X的分布而言，几乎不可能验证。因此，最不利噪声是否存在这一问题一直悬而未决。本文的关键贡献在于解决这一问题，证明对于任意 X的分布，最不利噪声都存在。证明依赖于一个关于条件期望方差的收敛结果：若（ X_n 给定 $X_n + Y_n$），且（X_n， Y_n）弱收敛，则该结果成立（见定理 2）。

在继续证明最不利噪声的存在性之前，我们先简要说明最不利噪声的一些实际意义。在某些应用科学场景中，观测到的随机信号 X 可能波动很大，因而难以分析。在这种情况下，我们可能希望在保留信号 X 主要结构的同时简化它。一种可行的方法是考虑随机变量

$$Q = E\left[X \middle| X + \hat{Y}_\epsilon\right],$$

ˆ其中 Y_ϵ 是最不利噪声。随机变量 Q 具有两个重要性质：(i) Q 的方差不大于 X 的方差，并且在满足 var$Y \leq \epsilon^2$的 Y上，使 var $E[X|X+Y]$ 达到最小值，这意味着它的波动性较小；(ii) 注意到，当 ϵ较小时，ˆ ˆ $X+Y_\epsilon$接近 X，因此随机变量 Q作为 $X+Y_\epsilon$的函数，保留了 X的结构。

现在我们将所考虑的问题形式化。为避免平凡情形，我们假设X 是非退化的，且为平方可积随机变量。不失一般性，我们可以假设 X 的均值为 0。由于将 Y 平移一个常数不会改变条件期望$E[X|X + Y]$ ，我们只关注以下这类随机变量 Y:

$$\mathcal{V}_\epsilon(X) := \{Y : Y \text{ independent of } X,\ E[Y] = 0 \text{ and } E[Y^2] \leq \epsilon^2\}.$$

我们定义

$$L(X, \epsilon) := \inf_{Y \in \mathcal{V}_\epsilon(X)} \text{var}\, E[X|X + Y].$$

ˆ 那么，最不利噪声 Y_ϵ是一个随机变量 Y，属于 $\mathcal{V}_\epsilon(X)$，且其方差 $E[X|X+Y]$达到 $L(X, \epsilon)$。

在第 2节中，我们证明最不利噪声的存在性。也就是说，我们证明存在一个ˆ Y_ϵ，当方差为 $E[X|X + Y]$ 时，使方差$Y \leq \epsilon^2$达到最小值。一般而言，要证明值函数 $f(x)$存在极小值点，通常采用以下策略：

(i) 构造一个序列 $\{x_n\}_{n=1}^\infty$，使得 $f(x_n)$ 收敛到 inf $f(x)$; $_x$

(ii) 找到 $\{x_n\}_{n=1}^\infty$ 的一个收敛子列，该子列收敛到 x_*;

(iii) 然后证明 $f(x_*) = \inf f(x)$，由此得出 x_*是x极小值点。

本文确实遵循上述策略。首先，我们考虑一个序列 $\{Y_n\}_{n=1}^{\infty}$，使得 $\mathrm{var} Y_n \le \epsilon^2$，并且 $E[X|X+Y_n]$ 的方差收敛至 $\inf_Y \mathrm{var}\ E[X|X+Y]$。其次，由于 $\{Y_n\}_{n=1}^{\infty}$是紧的，存在一个子序列弱收敛~到某个随机变量，记为 Y。接下来只需证明 $E[X|X+Y]$的方差恰好等于 $\inf_Y \mathrm{var}\, E[X|X+Y]$。这是本研究中数学上最具挑战性的任务。为此，我们建立了一个关于条件期望方差的收敛结果（见定理 2），其部分灵感来自 [2]的工作。

本文其余部分安排如下。在第 3 节中，我们证明最不利噪声的方差必定为 ϵ^2，从而在寻找最不利噪声时，可以缩小 Y 的类别。事实上，这一结果将 [3]中的三个刻画条件简化为两个刻画条件。在第 4 节中，我们进一步考察 $L(X, \epsilon)$ 的性质。我们证明它关于 ϵ 在 $[0, \infty)$ 上严格递减且右连续。因此，预测误差 $E[X - E[X|X+Y]]^2$的最大值会随着噪声 Y 的方差增大而严格增大。换言之，“噪声越大，预测效果越差。”这也彻底回答了 [3]第 2 页提出的问题 4。

2 最不利噪声的存在性

本节的主要结果是证明最不利噪声的存在性，我们将在定理 4 中给出。

我们先记录下文的引理 1。由于该引理是斯科罗霍德表示定理、法图引理以及 [1], 中定理 4.5.2 的直接推论，我们省略其证明。

引理 1. 设 $\{X_n\}_{n=1}^{\infty}$ 是一列弱收敛到某个随机变量 X 的随机变量。此外，假设

$$\sup_{1 \le n < \infty} E[X_n^2] < \infty.$$

Then

$$E[X^2] \le \liminf_{n\to\infty} E[X_n^2] \quad \textit{and} \quad \lim_{n\to\infty} E[X_n] = E[X].$$

下面我们将介绍定理 2 ，该定理给出了条件期望方差的一个收敛结果。它部分受到了 [2] 的工作启发。

定理 2. 设 $\{(X_n, Y_n)\}_{n=1}^{\infty}$ 是一个随机向量序列，弱收敛于某个随机向量 (X, Y)。如果

$$\sup_{1 \le n < \infty} E[X_n^2] < \infty,$$

then

$$\operatorname{var} E[X|X+Y] \leq \liminf_{n\to\infty} \operatorname{var} E[X_n|X_n+Y_n].$$

证明。 我们用 sup $E[X_n^2]$ 表示 M。根据连续映射定理， X_n 弱收敛于 X。应用引理1可得

$$E[X^2] \leq \liminf E[X_n^2] \leq M. \tag{1}$$

$$\lim_{n\to\infty} E[X_n] = E[X]. \tag{2}$$

根据条件期望的标准性质， $E[X_n|X_n+Y_n]$ 的均值为 $E[X_n]$。于是

$$\operatorname{var} E[X_n|X_n+Y_n] = E\left[(E[X_n|X_n+Y_n])^2\right] - (E[X_n])^2.$$

类似地，方差 $E[X|X+Y] = E\left[(E[X|X+Y])^2\right] - (E[X])^2$。结合（2），我们只需证明

$$E\left[(E[X|X+Y])^2\right] \leq \liminf_{n\to\infty} E\left[(E[X_n|X_n+Y_n])^2\right]. \tag{3}$$

由于(X_n, Y_n)弱收敛于(X, Y)，根据斯科罗霍德表示定理，我们可以在一个新的概率空间上构造(U_n, W_n)和(U, W)，使得$(U_n, W_n) \overset{d}{=} (X_n, Y_n)$、$(U, W) \overset{d}{=} (X, Y)$，并且$(U_n, W_n)$几乎必然收敛到$(U, W)$。由此立即可得

$$\sup_{1\leq n<\infty} E[U_n^2] = \sup_{1\leq n<\infty} E[X_n^2] = M,$$

and

$$E[U^2] = E[X^2] \leq M.$$

注意到， $E[X_n|X_n+Y_n]$ 和 $E[U_n|U_n+W_n]$ 的分布分别仅取决于(X_n, Y_n)和(U_n, W_n)的分布，由此可知

$$(X_n, Y_n) \overset{d}{=} (U_n, W_n)$$

即 d istribution 与 $E[X_n|X_n+Y_n]$ 的分布和 $E[U_n|U_n+W_n]$ 的分布一致。因此，

$$E\left[(E[X_n|X_n+Y_n])^2\right] = E\left[(E[U_n|U_n+W_n])^2\right].$$

类似地，

$$E\left[(E[X|X+Y])^2\right] = E\left[(E[U|U+W])^2\right].$$

因此，要证明（3），只需证明

$$E\left[(E[U|U+W])^2\right] \leq \liminf_{n\to\infty} E\left[(E[U_n|U_n+W_n])^2\right]. \tag{4}$$

我们接下来定义

$$T_n := E[U_n|U_n + W_n] \quad \text{and} \quad T := E[U|U + W]. \tag{5}$$

根据詹森不等式，对于所有 $n \in \mathbf{N}+$,

$$E[T_n^2] = E\left[(E[U_n|U_n + W_n])^2\right] \leq E\left[E[U_n^2|U_n + W_n]\right] = E[U_n^2] \leq M.$$

同理， $E[T^2] \leq M$。

证明接下来将运用以下引理，其证明见附录。

引理 3。对于 h 上的任意有界博雷尔函数 $\mathbb{R}$,

$$E[h(U + W)\, T] = \lim_{n\to\infty} E[h(U + W)\, T_n]. \tag{6}$$

现在说明引理 3 如何推出（4）。设 g 是定义在 $\mathbb{R}$上的博雷尔函数。根据条件期望的定义，$E[U|U + W]$ 可表示为 g（$U + W$）。对于 $k \in \mathbb{N}+$，令g_k ：$= g\, \mathbb{1}_{\{|g|\leq k\}}$。于是 g_k 是一个有界博雷尔函数。应用引理 3 可得

$$E[g_k(U + W)\, T] = \lim_{n\to\infty} E[g_k(U + W)\, T_n].$$

由于

$$E[g_k(U + W)\, T_n] \leq \left(E\left[(g_k(U + W))^2\right] + E[T_n^2]\right)/2,$$

令 $n \to \infty$ ，得到

$$E[g_k(U + W)\, T] \leq \frac{1}{2} \liminf_{n\to\infty} E[T_n^2] + \frac{1}{2} E\left[(g_k(U + W))^2\right]. \tag{7}$$

注意到

$$g(U + W) = E[U|U + W] = T$$

是平方可积的，由控制收敛定理可得

$$\lim_{k\to\infty} E\left[(g_k(U + W))^2\right] = E\left[(g(U + W))^2\right] = E[T^2]. \tag{8}$$

此外，注意到

$$|g_k(U + W)\, T| \leq |g(U + W)\, T| = T^2,$$

并回忆起 T^2 是可积的，随机变量族 $\{g_k(U+W)\ T\}$ 一致可积。由于 $g_k(U + W)\ T$ 几乎必然收敛到 $g(U + W)\ T$，我们有

$$\lim_{k\to\infty} E[g_k(U + W)\, T] = E[g(U + W)\, T] = E[T^2]. \tag{9}$$

最后，将 $k \to \infty$ 代入等式 (7) 的两边，并结合 (8) 和 (9) 中的结果，可得

$$E[T^2] \leq \liminf_n E[T_n^2],$$

这正是（4）中出现的内容。至此，证明完成。 □

有了定理 2 ，我们现在可以陈述本文关于最不利噪声存在性的关键定理。

定理 4. 假设 X 是一个零均值且具有有限二阶矩的非退化随机变量。那么存在一个极小值点 $Y \in \mathcal{V}_\epsilon(X)$，使得

$$\operatorname{var} E[X|X+Y] = \inf_{Z \in \mathcal{V}_\epsilon(X)} \operatorname{var} E[X|X+Z] = L(X,\epsilon).$$

因此，最不利噪声存在。

证明. 令 $Y_n \in \mathcal{V}_\epsilon(X)$ 为一个满足下述条件的随机变量

$$\operatorname{var} E[X|X+Y_n] \le L(X,\epsilon) + \frac{1}{n}.$$

由于 $E[X^2] < \infty$ 且 $\sup_n E[Y_n^2] \le \epsilon^2$，随机向量序列 $\{(X,Y_n)\}_{n=1}^{\infty}$ 是紧的。因此，存在 $\{(X,Y_n)\}$ 的一个子序列 $\{(X,Y_{n(k)})\}$，使得 $(X,Y_{n(k)})$ 弱收敛到某个随机向量，记为 $(\tilde{X},\tilde{Y})$。由连续映射定理可知，$X \overset{d}{=} \tilde{X}$ 和 $Y_{n(k)}$ 弱收敛到 $\tilde{Y}$。注意到 X 与 Y_n独立，可以很容易地验证 $\tilde{X}$ 与 $\tilde{Y}$独立。此外，将引理 1 应用于序列$\{Y_{n(k)}\}$，可得

$$E[\tilde{Y}^2] \le \liminf_{k\to\infty} E[Y_{n(k)}^2] \le \epsilon^2 \quad \text{and} \quad E[\tilde{Y}] = \lim_{k\to\infty} E[Y_{n(k)}] = 0.$$

现在，我们可以构造一个随机变量 Y，使得 Y 与 X 独立，且 Y 与 $\tilde{Y}$ 同分布。现在我们断言，Y 是所需的极小值点。下面证明这一断言。事实上，由于

$$E[Y] = E[\tilde{Y}] = 0$$

and

$$E[Y^2] = E[\tilde{Y}^2] \le \epsilon^2,$$

则 $Y \in \mathcal{V}_\epsilon(X)$。此外，根据 X 与 Y独立的假设可知 $(X,\ Y) \overset{d}{=} (\tilde{X},\ \tilde{Y})$。沿用定理 2 证明中的类似论证，可得

$$\operatorname{var} E[X|X+Y] = \operatorname{var} E[\tilde{X}|\tilde{X}+\tilde{Y}].$$

最后，应用定理 2可得

$$\operatorname{var} E[\tilde{X}|\tilde{X}+\tilde{Y}] \le \liminf_{k\to\infty} \operatorname{var} E[X_{n(k)}|X_{n(k)}+Y_{n(k)}] = L(X,\epsilon).$$

因此，$\operatorname{var} E[X|X+Y] \le L(X,\epsilon)$，这证明了 Y是所需的极小值点。 □

3 最不利噪声的方差

本节旨在计算最不利噪声的方差（定理7）。定理7 依赖于下文的命题5 和引理6 。命题 5 是条件期望的标准结果，因此我们省略证明。

命题 5. 假设 X 是一个平方可积随机变量。设 $\mathcal{F}$ 和 $\mathcal{G}$ 是两个 σ-代数，且满足 $\mathcal{G} \subset \mathcal{F}$。那么

$$E\left[(E[X|\mathcal{G}])^2\right] \leq E\left[(E[X|\mathcal{F}])^2\right],$$

当且仅当 $E[X|\mathcal{G}] = E[X|\mathcal{F}]$ 几乎必然成立时，等号成立。

现在我们引入引理 6，它依赖于命题 5。

引理 6. 假设 X 是一个均值为零且具有有限二阶矩的非退化随机变量， Y 是一个与 X独立的任意随机变量。设 Z 是一个与 $\sigma(X,Y)$独立的非退化高斯随机变量。那么

$$\operatorname{var} E[X|X+Y+Z] < \operatorname{var} E[X|X+Y].$$

证明. 首先，我们注意到 $E[X|X+Y]$ 和 $E[X|X+Y+Z]$ 的均值均为 $E[X] = 0$。因此，只需证明

$$E\left[(E[X|X+Y+Z])^2\right] < E\left[(E[X|X+Y])^2\right].$$

应用命题5可得

$$E\left[(E[X|X+Y+Z])^2\right] \leq E\left[(E[X|X+Y,Z])^2\right] = E\left[(E[X|X+Y])^2\right], \tag{10}$$

其中在最后一个等式中，我们用到了以下事实：

$$E[X|X+Y,Z] = E[X|X+Y],$$

这一结论成立，是因为 Z 与 σ（X， Y）相互独立。接下来，我们只需排除以下情况：

$$E\left[(E[X|X+Y+Z])^2\right] = E\left[(E[X|X+Y])^2\right]. \tag{11}$$

我们通过反证法进行证明。为导出矛盾，假设方程（11）成立。结合（10）和命题 5，可得

$$E[X|X+Y+Z] = E[X|X+Y,Z] = E[X|X+Y], \tag{12}$$

几乎必然成立。

设 f_1 和 f_2 是定义在 $\mathbb{R}$上的两个博雷尔函数。设 $f_1(X+Y+Z)$ 和 $f_2(X+Y)$ 分别是 $E[X|X+Y+Z]$ 和 $E[X|X+Y]$ 的版本。那么，式 (12) 中的等式意味着

$$P\left(f_1(X+Y+Z)=f_2(X+Y)\right)=1. \tag{13}$$

下文中，我们用 P_U 表示由随机变量 U在 $\mathbb{R}$ 上诱导的概率测度。由于 X、Y和 Z相互独立，式 (13) 等价于

$$\int_{\mathbb{R}}\int_{\mathbb{R}} P(f_1(x+y+Z)=f_2(x+y))\,P_X(dx)\,P_Y(dy)=1.$$

由于 $P(f_1(x+y+Z)=f_2(x+y))\le 1$，我们有

$$P(f_1(x+y+Z)=f_2(x+y))=1 \qquad P_X\otimes P_Y \text{ - a.s..}$$

于是我们可以选取 x_0、y_0，使得

$$P(f_1(x_0+y_0+Z)=f_2(x_0+y_0))=1,$$

这意味着 $f_1(x_0+y_0+Z)$ 几乎必然为常数。注意，由于 Z 是密度处处为正的绝对连续随机变量，f_1 相对于勒贝格测度几乎处处为常数。根据卷积的标准性质，$X+Y+Z$ 是绝对连续随机变量。因此，$f_1(X+Y+Z)$ 几乎必然为常数，也就是说，$E[X|X+Y+Z]$ 几乎必然为常数。注意，由于 $E[X|X+Y+Z]$ 的均值为 0，$E[X|X+Y+Z]$ 必定几乎处处为 0。于是

$$\begin{aligned} 0 &= E\left[E[X|X+Y+Z](X+Y+Z)\right]=E\left[X(X+Y+Z)\right] \\ &= E[X^2]+E[XY]+E[XZ]=E[X^2]+E[X]E[Y]+E[X]E[Z]=E[X^2], \end{aligned}$$

这意味着 $X=0$ 几乎必然。由于 X 是非退化的，因此得出了矛盾。证明到此结束。

□

有了上述引理，我们现在转而讨论最不利噪声的方差。

定理 7. 假设 X 是一个均值为零且具有有限二阶矩的非退化随机变量。对于任意满足 var $E[X|X+Y]$ 达到最小值 $L(X,\epsilon)$ 的 $Y\in\mathcal{V}_\epsilon(X)$，必有 $E[Y^2]=\epsilon^2$。

证明。 我们用反证法。为导出矛盾，假设 $E[Y^2]<\epsilon^2$。我们着手构造一个均值为零且方差足够小的高斯随机变量 Z，使得 Z与 $\sigma(X,Y)$ 独立，并且 $E[(Y+Z)^2]\le\epsilon^2$。不难验证 $Y+Z\in\mathcal{V}_\epsilon(X)$。然而，应用引理 6 可得

$$\operatorname{var} E[X|X+Y+Z]<\operatorname{var} E[X|X+Y].$$

然而，这与 Y 是极小值点这一事实直接矛盾。因此，根据反证法，$E[Y^2]=\epsilon^2$。

□

注 8. 定理7 告诉我们，为了寻找最不利噪声，只需关注满足 $Y \quad E[Y^2] = \epsilon^2$ 的随机变量。如果尝试用拉格朗日乘数法寻找最不利噪声，定理7 会将不等式约束转化为等式约束，从而大大简化问题。

4 $L(X, \epsilon)$ 的性质

本节旨在研究函数 $L(X, \epsilon)$ 的一些性质。引入命题9 是为了证明本节的关键结果（定理10），该结果指出，$L(X, \epsilon)$ 关于 ϵ 在 $[0, \infty)$ 上严格递减且右连续。

命题 9. 设 $\{X_n\}_{n=1}^{\infty}$ 是一列零均值随机变量，弱收敛于某个零均值随机变量 X。设 $\{\epsilon_n\}_{n=1}^{\infty}$ 是一列非负实数，收敛于非负实数 ϵ。如果

$$\sup_{1 \le n < \infty} E[X_n^2] < \infty,$$

then

$$L(X, \epsilon) \le \liminf_{n \to \infty} L(X_n, \epsilon_n).$$

证明。根据下极限的性质，存在 $\{n\}_{n=1}^{\infty}$ 的一个子序列 $\{n(k)\}_{k=1}^{\infty}$，使得

$$\liminf_{n \to \infty} L(X_n, \epsilon_n) = \lim_{k \to \infty} L(X_{n(k)}, \epsilon_{n(k)}).$$

因此，只要 $L(X_{n(k)}, \epsilon_{n(k)})$ 的极限存在，我们只需证明

$$L(X, \epsilon) \le \liminf_{k \to \infty} L(X_{n(k)}, \epsilon_{n(k)}).$$

为简化记号，在下文中，我们仍将子序列 $\{n(k)\}_{k=1}^{\infty}$记为 $\{n\}_{n=1}^{\infty}$。我们还假设 $L(X_n, \epsilon_n)$ 的极限存在。

对于每个 n，根据定理 4，存在一个 $Y_n \in \mathcal{V}_{\epsilon_n}(X_n)$，使得

$$\mathrm{var} E[X_n | X_n + Y_n] = L(X_n, \epsilon_n).$$

注意到 $E[Y_n^2] \le \epsilon_n^2$ 且 ϵ_n 收敛于 ϵ，我们有 $\sup_n E[Y_n^2] < \infty$。结合 $\sup_n E[X_n^2] < \infty$这一事实，随机向量族 $\{(X_n, Y_n)\}_{n=1}^{\infty}$ 是紧的。因此，存在 $\{(X_{n(k)}, Y_{n(k)})\}$ 这一子序列，它是 $\{(X_n, Y_n)\}$ 的子序列，并且 $(X_{n(k)}, Y_{n(k)})$ 收敛~ ~弱收敛于某个随机向量，记为 (X, Y)。由连续映射可得，$X_{n(k)}$弱收敛于 $\tilde{X}$，且 $Y_{n(k)}$ 弱收敛于 $\tilde{Y}$。于是，$X \overset{d}{=} \tilde{X}$，因为 $X_{n(k)}$ 也弱收敛于 X。由引理 1 可得 $E[\tilde{Y}] = \lim_k E[Y_{n(k)}] = 0$ 以及 $E[\tilde{Y}^2] \le$

下极限$_k$ $E[Y^2_{n(k)}] \le \epsilon^2$。此外，由于 $X_{n(k)}$ 与 $Y_{n(k)}$独立，我们还有 $\tilde{X}$与 Y独立。应用定理 2 可得

$$\begin{aligned} & \operatorname{var}E[\tilde{X}|\tilde{X}+\tilde{Y}] \le \liminf_{k\to\infty} \operatorname{var}E[X_{n(k)}|X_{n(k)}+Y_{n(k)}] \\ = & \liminf_{k\to\infty} L(X_{n(k)},\epsilon_{n(k)}) = \liminf_{n\to\infty} L(X_n,\epsilon_n), \end{aligned} \tag{14}$$

其中，最后一个等式由我们的假设得出，即 $L(X_n,\epsilon_n)$ 的极限存在。

存在一个随机变量 Y，使得 Y 与 X 独立，且 $Y \overset{d}{=} Y$。因此，$E[Y]=E[\tilde{Y}]=0$ 和 $E[Y^2]=E[\tilde{Y}^2]\le\epsilon^2$，这意味着 $Y\in\mathcal{V}_\epsilon(X)$。由独立性可得，$(X,Y) \overset{d}{=} (\tilde{X},\tilde{Y})$，于是，

$$\operatorname{var}E[X|X+Y] = \operatorname{var}E[\tilde{X}|\tilde{X}+\tilde{Y}], \tag{15}$$

根据证明定理 2 时的类似论证。结合 (14) 和 (15)，可得

$$L(X,\epsilon) \le \operatorname{var}E[X|X+Y] \le \liminf_{n\to\infty} L(X_n,\epsilon_n).$$

证明完毕。 □

定理 10. 设 X 是一个均值为零且具有有限二阶矩的非退化随机变量。那么，固定 X 时，$L(X,\epsilon)$ 关于 ϵ 在 $[0,\infty)$ 上是严格递减且右连续的函数。

证明。 我们首先证明 $L(X,\epsilon)$ 关于 ϵ严格递减。考虑$0\le\epsilon_1<\epsilon_2$。根据定理 4，存在 $Y_1\in\mathcal{V}_{\epsilon_1}(X)$，使得

$$\operatorname{var} E[X|X+Y_1] = L(X,\epsilon_1).$$

我们继续构造一个均值为零、方差为 $\epsilon_2^2-\epsilon_1^2$的高斯随机变量，使得 Z与 $\sigma(X,Y_1)$ 独立。容易验证 $Y_1+Z\in\mathcal{V}_{\epsilon_2}(X)$。于是根据引理 6，

$$L(X,\epsilon_2) \le \operatorname{var} E[X|X+Y_1+Z] < \operatorname{var} E[X|X+Y_1] = L(X,\epsilon_1),$$

这意味着 $L(X,\epsilon)$ 严格递减。

现在我们转而证明右连续性。考虑每个 $\epsilon_0\in[0,\infty)$。对于 $\epsilon>\epsilon_0$，我们有$L(X,\epsilon) < L(X,\epsilon_0)$。因此

$$\limsup_{\epsilon\downarrow\epsilon_0} L(X,\epsilon) \le L(X,\epsilon_0). \tag{16}$$

应用命题 9可得

$$L(X,\epsilon_0) \le \liminf_{\epsilon\downarrow\epsilon_0} L(X,\epsilon). \tag{17}$$

最后，结合（16）和（17），即可完成证明。 □

注 11. 定理 *10* 表明，预测误差的最大值 $E[X - E[X|X + Y]]^2$ 会随着噪声 Y 的方差增大而严格递增。换言之，“噪声越大，预测效果越差。”这就明确回答了第 2 页提出的问题 4 *[3]*。

致谢：感谢我的前博士导师 Philip A. Ernst 教授就本文进行了许多富有帮助且见解深刻的讨论。

参考文献

[1] K.L. Chung。《概率论教程》。Academic Press，2001年。

[2] I. Crimaldi 和 L. Pratelli。条件期望的收敛结果。Bernoulli，11(4)：737–745，2005。[3] P.A. Ernst、A.M. Kagan 和 L.C.G. Rogers。最不利噪声。Electronic Communications in Probability，27：1–11，2022。[4] L.C.G. Rogers 和 D. Williams。扩散过程、马尔可夫过程与鞅：第 1 卷，基础。剑桥大学出版社，2000 年。

5 附录

本附录包含引理 3 的证明。

引理的证明 3。 我们首先断言，类 $\mathcal{H}$ 上的有界实值博雷尔函数 h，满足 (6) 的条件，则必定具有以下性质 (A.1)-(A.3)：

(A.1) $\mathcal{H}$ 是一个包含常值函数的向量空间；

(A.2) $\mathcal{H}$ 在一致收敛下封闭；

(A.3) 对于 $\mathcal{H}$ 中的一致有界非负函数序列 $\{h_k\}$，其中 $\forall k, h_k \leq h_{k+1}$，且 $\forall s$, $h_k(s) \to h(s)$，有 $h \in \mathcal{H}$。

显然， $\mathcal{H}$ 是一个向量空间。此外，对于任意常值函数 $c(x) \equiv c$，根据 (2)，有

$$\lim_{n\to\infty} E[c(U+W)\,T_n] = \lim_{n\to\infty} c\,E[T_n] = \lim_{n\to\infty} c\,E[X_n] = c\,E[X] = c\,E[T] = E[c(U+W)\,T].$$

因此，性质 (A.1) 成立。

为验证性质 (A.2)，假设 $h_k \in \mathcal{H}$ 一致收敛到 h。那么

$$\begin{aligned}
& |E[h(U+W)\,T] - E[h(U+W)\,T_n]| \\
\leq\ & |E\left[(h(U+W) - h_k(U+W))\ T\right]| + |E\left[(h(U+W) - h_k(U+W))\ T_n\right]| \\
& + |E[h_k(U+W)\,T] - E[h_k(U+W)\,T_n]| \\
\leq\ & \left\{E\left[(h(U+W) - h_k(U+W))^2\right]\right\}^{1/2} \cdot \left\{E[T^2]\right\}^{1/2} \\
& + \left\{E\left[(h(U+W) - h_k(U+W))^2\right]\right\}^{1/2} \cdot \left\{E[T_n^2]\right\}^{1/2} \\
& + |E[h_k(U+W)\,T] - E[h_k(U+W)\,T_n]| \\
\leq\ & 2M^{1/2} \left\{E\left[(h(U+W) - h_k(U+W))^2\right]\right\}^{1/2} \\
& + |E[h_k(U+W)\,T] - E[h_k(U+W)\,T_n]|,
\end{aligned}$$

其中第二个不等式由 H 推出¨older 不等式得出，而在最后一个不等式中，我们应用了 sup 的事实$_n$ $E[T_n^2] \leq M$且 $E[T^2] \leq M$。令 n趋于 ∞，并注意到 (6) 对 h_k成立，我们有

$$\limsup_{n\to\infty} |E[h(U+W)T] - E[h(U+W)T_n]| \leq 2M^{1/2} \left\{E\left[(h(U+W) - h_k(U+W))^2\right]\right\}^{1/2}. \tag{18}$$

令 k 趋于 ∞，并回想起 h_k 一致收敛于 h这一事实，我们得出结论： $h \in \mathcal{H}$。

现在我们转向 (A.3)。假设 $h_k \uparrow h$是有界的非负函数。显然， $h_k(U+W) - h(U+W)$逐点收敛于 0，且一致有界。根据控制收敛定理，我们有

$$\lim_{n\to\infty} E\left[(h(U+W) - h_k(U+W))^2\right] = 0 \tag{19}$$

用与证明（A.2）类似的论证，可再次得到（18）。结合（19），我们得出 $h \in \mathcal{H}$。

做好上述准备后，我们来证明（6）对每个有界博雷尔函数 h都成立。根据单调类定理（参见[4, 第91]页），只需证明（6）对每个定义在 $\mathbb{R}$上的有界连续函数 h 成立。回顾一下，根据构造，（U_n， W_n）几乎必然收敛到（U， W）。由于 h 是连续的，我们有 h（U_n+W_n）$\to h$（$U+W$）几乎必然成立。由于 h有界，根据控制收敛定理，我们有

$$\lim_{n\to\infty} E\left[(h(U_n+W_n) - h(U+W))^2\right] = 0. \tag{20}$$

注意到

$$\sup_{1\le n<\infty} E\left[(U_n h(U_n+W_n))^2\right] \le \|h\|^2 \cdot \sup_{1\le n<\infty} E[U_n^2] = M\|h\|^2,$$

其中 $\|h\| := \sup_x |h(x)|$。那么，随机变量族 $\{U_n h(U_n+W_n)\}_{n=1}^{\infty}$ 一致可积。此外， $U_n h(U_n+W_n)$ 几乎必然收敛到 $Uh(U+W)$，因此

$$\lim_{n\to\infty} E[U_n h(U_n+W_n)] = E[Uh(U+W)]. \tag{21}$$

回顾（ T_n 和 T ）中给出的定义，并运用条件期望的标准性质，我们有

$$E[h(U_n+W_n)\,T_n] = E\left[h(U_n+W_n)E[U_n|U_n+W_n]\right] = E[U_n h(U_n+W_n)],$$

and

$$E[h(U+W)\,T] = E\left[h(U+W)E[U|U+W]\right] = E[Uh(U+W)].$$

将上述两个表达式代入（21）可得

$$\lim_{n\to\infty} E[h(U_n+W_n)\,T_n] = E[h(U+W)\,T]. \tag{22}$$

根据 H¨older 不等式，我们有

$$\begin{aligned} & |E[h(U_n+W_n)\,T_n] - E[h(U+W)\,T_n]| \\ = \; & |E[(h(U_n+W_n) - h(U+W))\,T_n]| \\ \le \; & \left\{E\left[(h(U_n+W_n) - h(U+W))^2\right]\right\}^{1/2} \cdot \left\{E[T_n^2]\right\}^{1/2} \end{aligned}$$

$$\leq \quad M^{1/2} \left\{ E\left[(h(U_n + W_n) - h(U + W))^2 \right] \right\}^{1/2}.$$

结合（20），可得

$$\lim_{n \to \infty} |E[h(U_n + W_n)\, T_n] - E[h(U + W)\, T_n]| = 0 \tag{23}$$

结合（22）和（23），可得（6）中的等式。 □