7
Entropia da Transformação de Gauss

Neste capítulo calcularemos a Entropia da Transformação de Gauss (Teorema 7.1). Para isso, usaremos dois resultados importantes, o Teorema de Shannon-McMillan-Breiman e o Teorema de Kolmogorov-Sinai (para ver a prova destes teorema veja as referências (4), (6), (8) e (9)).

Teorema 7.1 *A entropia $h(T)$ da transformação de Gauss T em $[0,1)$ com respeito a medida de Gauss é*

$$h(T) = \frac{\pi^2}{6\log 2}.$$

Antes de provar este teorema introduziremos alguns conceitos.

Considere a quadrúpla $(X, \mathcal{F}, \nu, S)$, onde X é um conjunto não vazio, $\mathcal{F}$ é uma σ-álgebra definida em X, ν é uma medida de probabilidade associada à σ-álgebra $\mathcal{F}$, e S é uma transformação mensurável ergódica e sobrejetora. Seja $\alpha = \{A_i\}_{i\in I}$ uma partição enumerável ou finita de X. Sabemos que $S^{-1}(\alpha) = \{S^{-1}(A_i)\}_{i\in I}$ também é uma partição de X. Assim, indutivamente, $S^{-(n)}(\alpha) = \{S^{-(n)}(A_i)\}_{i\in I}$ é uma partição de X.

Dadas duas partições de X, $\beta = \{B_j\}_{j\in J}$ e $\gamma = \{C_j\}_{j\in H}$, denotamos por $\beta \vee \gamma$ a nova partição de X definida como

$$\beta \vee \gamma = \{B_i \cap C_j; \quad B_i \in \beta \text{ e } C_j \in \gamma\}.$$

Temos portanto definida, para cada par de números inteiros k e n, $k < n$, a partição

$$\bigvee_{i=k}^{n} S^{-i}(\alpha) = S^{-k}(\alpha) \vee S^{-k-1}(\alpha) \vee \ldots \vee S^{-(n)}(\alpha).$$

Finalmente, denotamos por $\sigma\left(\bigvee_{i=-\infty}^{\infty} S^{-i}(\alpha)\right)$ a menor σ-álgebra contendo as partições

$$\bigvee_{i=m}^{l} S^{-i}(\alpha) \quad \text{e} \quad \bigvee_{i=-l}^{-m} S^{-i}(\alpha).$$

Se $\sigma\left(\bigvee_{i=-\infty}^{\infty} S^{-i}(\alpha)\right) = \mathcal{F}$, dizemos que α é um *gerador* com respeito a S.

Agora estamos prontos para definir a *Entropia de uma partição* α de X e, a partir dela, a *Entropia da transformação* S.

Definição 7.2 *Considere um conjunto X, uma medida de probabilidade ν definida em X, uma transformação mensurável $S\colon X \to X$ que preserva a medida, e uma partição enumerável $\alpha = \{A_i\}_{i\in I}$ de X formada por Borelianos. Definimos a* entropia de uma partição α *por*

$$H(\alpha) = -\sum_{i\in I} \nu(A_i)\,\log(\nu(A_i)).$$

Então, a entropia com respeito a partição α *da transformação S é dada por*

$$h(\alpha, S) = \lim_{n\to\infty} \frac{1}{n} H\left(\bigvee_{i=0}^{n-1} S^{-i}(\alpha)\right).$$

E, finalmente, a entropia de S *é*

$$h(S) = \sup_{\alpha} h(\alpha, S).$$

onde o supremo é considerado no conjunto das partições enumeráveis de X.

Observamos que o limite na equação de $h(\alpha, S)$ existe. Para provar essa observação precisamos de duas propriedades de entropia e de uma proposição que enunciaremos a seguir:

Propriedade (I): Para $\mathcal{P}_1$ e $\mathcal{P}_2$ partições finitas de um espaço de probabilidade X, temos

$$H(\mathcal{P}_1 \vee \mathcal{P}_2) \leq H(\mathcal{P}_1) + H(\mathcal{P}_2).$$

Prova: Sejam $\mathcal{P}_1 = \{A_1, \ldots, A_n\}$ e $\mathcal{P}_2 = \{B_1, \ldots, B_m\}$ as duas partições do espaço de probabilidade X. Para provar a Propriedade (I), precisamos em primeiro lugar definir a *entropia condicional* de $\mathcal{P}_2$ com respeito à $\mathcal{P}_1$ como

$$H(\mathcal{P}_2|\mathcal{P}_1) = -\sum_{i\in I, j\in J} \nu(B_i \cap A_j)\,\log\left(\frac{\nu(B_i \cap A_j)}{\nu(A_j)}\right),$$

onde $I = \{1, \ldots, m\}$ e $J = \{1, \ldots, n\}$.

Temos agora duas afirmações:

Afirmação 7.3

$$H(\mathcal{P}_1 \vee \mathcal{P}_2) = H(\mathcal{P}_1) + H(\mathcal{P}_2|\mathcal{P}_1).$$

Afirmação 7.4

$$H(\mathcal{P}_2|\mathcal{P}_1) \leq H(\mathcal{P}_2).$$

Posporemos a prova destas afirmações e agora veremos que estas afirmações implicam a Propriedade (I).

Pela Afirmação 7.3 e pela Afirmação 7.4, obtemos que

$$H(\mathcal{P}_1 \vee \mathcal{P}_2) = H(\mathcal{P}_1) + H(\mathcal{P}_2|\mathcal{P}_1) \leq H(\mathcal{P}_1) + H(\mathcal{P}_2).$$

Para concluir a prova da Propriedade (I) precisamos provar a Afirmação 7.3 e a Afirmação 7.4.

Prova da Afirmaçaõ 7.3: Temos que

$$\begin{aligned} H(\mathcal{P}_1 \vee \mathcal{P}_2) &= -\sum_{i\in I, j\in J} \nu(A_j \cap B_i)\,\log(\nu(A_j \cap B_i)) = \\ &= -\sum_{i\in I, j\in J} \nu(A_j \cap B_i)\,\log\left(\frac{\nu(A_j \cap B_i)\,\nu(A_j)}{\nu(A_j)}\right) = \\ &= -\sum_{i\in I, j\in J} \nu(A_j \cap B_i)\,\log\left(\nu(A_j)\right) - \\ &\quad -\sum_{i\in I, j\in J} \nu(A_j \cap B_i)\,\log\left(\frac{\nu(A_j \cap B_i)}{\nu(A_j)}\right). \end{aligned}$$

Como $\mathcal{P}_1$ e $\mathcal{P}_2$ são partições de um espaço de probabilidade, obtemos

$$\begin{aligned}
H(\mathcal{P}_1 \vee \mathcal{P}_2) &= -\sum_{i\in I, j\in J} \nu(A_j \cap B_i)\, \log\left(\nu(A_j)\right) - \\
&\quad - \sum_{i\in I, j\in J} \nu(A_j \cap B_i)\, \log\left(\frac{\nu(A_j \cap B_i)}{\nu(A_j)}\right) = \\
&= -\sum_{j\in J} \nu(A_j)\, \log\left(\nu(A_j)\right) - \\
&\quad - \sum_{i\in I, j\in J} \nu(A_j \cap B_i)\, \log\left(\frac{\nu(A_j \cap B_i)}{\nu(A_j)}\right) = \\
&= H(\mathcal{P}_1) + H(\mathcal{P}_2|\mathcal{P}_1).
\end{aligned}$$

Isto termina a prova da afirmação.

Prova da Afirmaçaõ 7.4: Temos que

$$\begin{aligned}
H(\mathcal{P}_2|\mathcal{P}_1) &= -\sum_{i\in I, j\in J} \nu(A_j \cap B_i)\, \log\left(\frac{\nu(A_j \cap B_i)}{\nu(A_j)}\right) \leq \\
&\leq -\sum_{i\in I, j\in J} \nu(A_j \cap B_i)\, \log\left(\nu(A_j \cap B_i)\right) = \\
&= -\sum_{i\in I} \nu(B_i)\, \log\left(\nu(B_i)\right) = H(\mathcal{P}_2).
\end{aligned}$$

Completamos assim a prova da afirmação.

Desta forma, concluímos a demonstração da Propriedade (I). $\square$

Propriedade (II): Se a medida de probabilidade é S-invariante, então para $\mathcal{P}_1$ uma partição finita de um espaço de probabilidade X temos

$$H(S^{-1}\mathcal{P}_1) = H(\mathcal{P}_1).$$

Prova: Tome $\mathcal{P}_1 = \{A_1, \ldots, A_n\}$ uma partição de um espaço de probabilidade e seja J o conjunto $J = \{1, \ldots, n\}$.

Para provar a Propriedade (II), note que pela invariância de S,

$$H(S^{-1}\mathcal{P}_1) = -\sum_{j\in J}\nu(S^{-1}(A_j))\,\log\big(\nu(S^{-1}(A_j))\big) =$$

$$= -\sum_{j\in J}\nu(A_j)\,\log(\nu(A_j)) = H(\mathcal{P}_1).$$

Isto termina a prova da Propriedade (II). □

Observe que $H(S^{-m}(\mathcal{P}_1)) = H(\mathcal{P}_1)$. Indutivamente, suponha que para toda partição finita $\mathcal{P}_1 = \{A_1, \ldots, A_n\}$ e para todo $1 \leq j < m$,

$$H(S^{-(j-1)}(\mathcal{P}_1)) = H(\mathcal{P}_1).$$

Então, usando a hipótese de indução e a Propriedade (II),

$$H(S^{-m}(\mathcal{P}_1)) = H(S^{-(m-1)}(S^{-1}(\mathcal{P}_1))) = H(S^{-1}(\mathcal{P}_1)) = H(\mathcal{P}_1).$$

Proposição 7.5 *Seja $(a_n)_{n\geq 1}$ uma seqüência de números reais subaditiva, isto é, $a_{n+m} \leq a_n + a_m$. Então, o limite*

$$\lim_{n\to\infty}\frac{a_n}{n}$$

existe.

Prova: Seja $L = \liminf_{n\to\infty}\frac{a_n}{n}$. Dado $\epsilon > 0$, seja n_0 tal que $\frac{a_{n_0}}{n_0} \leq L + \epsilon$.

Se $n > n_0$, então podemos escrever

$$n = n_0\,p + q, \quad \text{onde } p, q \in \mathbb{Z} \text{ com } 0 < q \leq n_0 \text{ e } p \geq 1.$$

Então, como $a_{n+m} \leq a_n + a_m$,

$$\frac{a_n}{n} = \frac{a_{n_0\,p+q}}{n_0\,p+q} \leq \frac{a_{n_0\,p} + a_q}{n_0\,p+q} = \frac{a_{n_0\,p}}{n_0\,p+q} + \frac{a_q}{n_0\,p+q} \leq$$

$$\leq \frac{p\,a_{n_0}}{n_0\,p} + \frac{a_q}{n_0\,p} = \frac{a_{n_0}}{n_0} + \frac{a_q}{n_0\,p}.$$

Tomando $\limsup_{n\to\infty}\frac{a_q}{n_0\,p} \leq L$, e como $L = \liminf_{n\to\infty}\frac{a_n}{n}$, obtemos que,

$$L \leq \frac{a_n}{n} \leq L + \epsilon + L = 2\,l + \epsilon.$$

Portanto, como $\epsilon > 0$ é arbitrário, temos o resultado. $\square$

Agora estamos prontos para provar que o limite na equação de $h(\alpha, S)$ na Definição 7.2 existe.

Seja

$$H_v = H\left(\bigvee_{i=0}^{v-1} S^{-i}(\alpha)\right).$$

Usando as Propriedades (I) e (II), temos que

$$\begin{aligned}
H_{m+n} &= H\left(\alpha \vee \ldots \vee S^{-(m-1)}(\alpha) \vee S^{-m}(\alpha) \vee \cdots \vee S^{-(m+n-1)}(\alpha)\right) \leq \\
&\leq H\left(\alpha \vee \ldots \vee S^{-(m-1)}(\alpha)\right) + H\left(S^{-m}(\alpha) \vee \cdots \vee S^{-(m+n-1)}(\alpha)\right) = \\
&= H\left(\alpha \vee \ldots \vee S^{-(m-1)}(\alpha)\right) + H\left(S^{-m}\left(\alpha \vee \cdots \vee S^{-(n-1)}(\alpha)\right)\right) = \\
&= H\left(\alpha \vee \ldots \vee S^{-(m-1)}(\alpha)\right) + H\left(\alpha \vee \cdots \vee S^{-(n-1)}(\alpha)\right) \\
&= H_m + H_n.
\end{aligned}$$

Portanto, H_n é uma seqüência subaditiva.

Logo, aplicado a Proposição 7.5, obtemos que o limite

$$\lim_{n\to\infty} \frac{H_n}{n}$$

existe, isto termina a prova da observação.

Dois importantes teoremas que serão utilizados para calcular a entropia da Transformação de Gauss são os seguintes.

Teorema 7.6 (Shannon-McMillan-Breiman) *Seja $(X, \mathcal{F}, \nu, S)$ uma quadrúpla como acima, e seja α uma partição finita ou enumerável de X com $H(\alpha) < \infty$. Então para quase todo $x \in X$,*

$$\lim_{n\to\infty} -\frac{1}{n} \log(\nu(A_n(x))) = h(\alpha, S),$$

onde $A_n(x)$ é o elemento da partição $\bigvee_{i=0}^{n-1} S^{-i}(\alpha)$ que contém x.

Teorema 7.7 (Teorema de Kolmogorov-Sinai) *Se α é um gerador finito ou enumerável para a transformação S com $H(\alpha) < \infty$, então*

$$h(S) = h(\alpha, S).$$

Agora estamos em condições de calcular a entropia da Transformação de Gauss T com respeito á medida ergódica de Gauss. A partir de agora consideraremos $X = [0, 1)$, $\mathcal{B}$ a σ-álgebra dos Borelianos, $\nu = \mu$ a medida de Gauss, e $S = T$ a Transformação de Gauss. Veremos que neste caso, a partição dos cilindros $\alpha = \{I_1, I_2, \ldots, I_n, \ldots\}$ introduzidos na Seção 2.3.4 fornece um gerador enumerável da σ-álgebra $\mathcal{B}$ que verifica $H(\alpha) < \infty$. Isto nos permitirá usar o Teorema 7.7 para calcular a entropia de T com respeito a μ considerando exclusivamente a partição α.

Para ver que α é um gerador da σ-álgebra $\mathcal{B}$, para cada $n \geq 1$ escrevemos

$$I_{[n]} = I_{i_1,\ldots,i_n} = I_{i_1} \cap T^{-1}(I_{i_2}) \cap \ldots \cap T^{-(n-1)}(I_{i_n}).$$

Estes cilindros pertencem à partição $\bigvee_{i=1}^{n-1} T^{-i}(\alpha)$ e como os cilindros $I_{[n]}$ geram a σ-ágebra dos borelianos $\mathcal{B}$, temos que a partição α é um gerador com respeito a transformação de Gauss T. A seguir veremos que a entropia da partição é finita.

Lema 7.8 *$H(\alpha) < \infty$.*

Prova: Em primeiro lugar, pela Observação 5.6

$$\mu(I_n) \leq \frac{1}{\log 2}\,\lambda(I_n) = \frac{1}{\log 2}\,\frac{1}{n(n+1)} < \frac{1}{n^2\,\log 2}.$$

Portanto, como

$$\begin{aligned} H(\alpha) &= -\sum_{n=1}^{\infty}\mu(I_n)\,\log(\mu(I_n)) < -\sum_{n=1}^{\infty}\frac{1}{n^2\,\log 2}\log\left(\frac{1}{n^2\,\log 2}\right) < \\ &< -\sum_{n=1}^{\infty}\frac{1}{n^2\,\log 2}\,\log\left(\frac{1}{n^2}\right) = -\sum_{n=1}^{\infty}\frac{1}{n^2\,\log 2}\,(-2)\,\log(n). \end{aligned}$$

é suficiente verificar que a soma

$$\sum_{n=1}^{\infty}\frac{2}{n^2\,\log 2}\,\log(n)$$

é convergente. Isto decorre do fato da integral

$$\int_1^{\infty}\frac{\log(x)}{x^2}\,dx$$

ser convergente (é suficiente integrar por partes). □

Portanto, estamos em condições de aplicar o Teorema 7.7 obtendo que

$$h(T) = h(\alpha, T).$$

Então agora precisamos nos preocupar apenas em calcular $h(\alpha, T)$. Para isso usaremos o Teorema 7.6 e o seguinte resultado:

Lema 7.9

$$\lim_{n\to\infty} \frac{\log(\lambda(I_{[n]}))}{\log(\mu(I_{[n]}))} = 1.$$

Prova: Pela Observação 5.6,

$$\frac{1}{2\log 2}\,\lambda(I_{[n]}) < \mu(I_{[n]}) \leq \frac{1}{\log 2}\,\lambda(I_{[n]}).$$

Obtemos então que

$$(\log 2)\,\mu(I_{[n]}) \leq \lambda(I_{[n]}) < (2\log 2)\,\mu(I_{[n]}).$$

Tomando logaritmos e usando as suas propriedades, temos que

$$\begin{aligned} \log(\log 2) + \log\left(\mu(I_{[n]})\right) &= \log\left(\log 2\,\mu(I_{[n]})\right) \\ &\leq \log(\lambda(I_{[n]})) \\ &< \log\left(2\,\log 2\,\mu(I_{[n]})\right) = \\ &= \log(2\,\log 2) + \log\left(\mu(I_{[n]})\right). \end{aligned}$$

Dividindo por $\log\left(\mu(I_{[n]})\right)$,

$$\frac{\log(\log 2) + \log\left(\mu(I_{[n]})\right)}{\log\left(\mu(I_{[n]})\right)} \leq \frac{\log(\lambda(I_{[n]}))}{\log\left(\mu(I_{[n]})\right)} < \frac{\log(2\,\log 2) + \log\left(\mu(I_{[n]})\right)}{\log\left(\mu(I_{[n]})\right)}.$$

Lembramos que $\lambda(I_{[n]}) \to 0$ quando $n \to \infty$. Portanto, $\mu(I_{[n]}) \to 0$ quando $n \to \infty$. Finalmente, como $\log(w) \to \infty$ quando $w \to 0$, temos que

$$\frac{\log(\lambda(I_{[n]}))}{\log\left(\mu(I_{[n]})\right)} \to 1$$

quando $n \to \infty$, terminando assim a prova do lema. $\square$

Veremos agora como através do Lema 7.9 fica fácil calcular a entropia da transformação de Gauss T usando o Teorema 7.6. Pelo Lema 7.9 temos que

$$\lim_{n\to\infty} -\frac{1}{n}\log(\mu(I_{[n]})) = \lim_{n\to\infty} -\frac{1}{n}\log(\lambda(I_{[n]})).$$

Então, pelo Teorema 7.6, existe um conjunto de medida total Y de $[0,1)$ tal que se $x \in Y$ e $I_{[n]}(x)$ é o cilindro de ordem n que contém x, então

$$h(\alpha, T) = \lim_{n\to\infty} -\frac{1}{n}\log(\lambda(I_{[n]}(x))).$$

Esta é a expressão que iremos calcular agora.

Escrevemos $I_{[n]}(x) = I_{[n]}$. Pela equação (5-2), temos

$$\begin{aligned}\log(\lambda(I_{[n]})) &= \log\left(\frac{1}{q_n\,(q_n+q_{n-1})}\right) = -\log\left(q_n\,(q_n+q_{n-1})\right)\\ &= -\log(q_n) - \log(q_n+q_{n-1}).\end{aligned}$$

Então, como $q_n \leq q_n + q_{n-1}$ temos que $-\log(q_n+q_{n-1}) \leq -\log(q_n)$ e como $2\,q_n = q_n + q_n \geq q_n + q_{n-1}$ temos que $-\log 2 - \log(q_n) \leq -\log(q_n+q_{n-1})$. Portanto,

$$\begin{aligned}-\log 2 - 2\log(q_n) &= -\log 2 - \log(q_n) - \log(q_n) <\\ &< -\log(q_n+q_{n-1}) - \log(q_n) = \log(\lambda(I_{[n]})) <\\ &< -\log(q_n) - \log(q_n) = -2\log(q_n).\end{aligned}$$

Dividindo por n,

$$-\frac{1}{n}\log 2 - 2\,\frac{1}{n}\log(q_n) < \frac{1}{n}\log(\lambda(I_{[n]})) < -2\,\frac{1}{n}\log(q_n).$$

Usando a Proposição 5.14 item (III) obtemos que, para quase todo x

$$h(\alpha, T) = \lim_{n\to\infty} -\frac{1}{n}\log(\lambda(I_{[n]})) = \frac{\pi^2}{6\log 2}.$$

Concluímos assim a prova do Teorema 7.1.