Entropia i jej własności

Jacek Jurkowski, Fizyka Statystyczna

Instytut Fizyki

2015

Entropia termodynamiczna

Entropia termodynamiczna

- $dS = \frac{dQ}{T}$ dla procesów odwracalnych

Entropia termodynamiczna

- $dS = \frac{dQ}{T}$ dla procesów odwracalnych
- i wówczas
$$\Delta S = \int_{A \to B} \frac{dQ}{T}$$
gdzie powyższa całka nie zależy od procesu $A \to B$, a tylko od stanu początkowego A i końcowego B

Entropia termodynamiczna

- $dS = \frac{dQ}{T}$ dla procesów odwracalnych
- i wówczas

$$\Delta S = \int_{A \rightarrow B} \frac{dQ}{T}$$

gdzie powyższa całka nie zależy od procesu $A \rightarrow B$, a tylko od stanu początkowego A i końcowego B

- Ogólnie $dS \geqslant \frac{dQ}{T}$, bowiem ma miejsce produkcja entropii $d_i S$,

$$dS = \frac{dQ}{T} + d_i S$$

Entropia statystyczna

Entropia statystyczna

- Niech $P = (p_1, \ldots, p_N)$ będzie prawdopodobieństwem (N może być nieskończone!)

Entropia statystyczna

- Niech $P = (p_1, \ldots, p_N)$ będzie prawdopodobieństwem (N może być nieskończone!)

Definicja

Entropię definiuje się jako

$$S(P) \;=\; -k \sum_{i=1}^{N} p_i \ln p_i \quad \text{albo} \quad S(P) \;=\; -\sum_{i=1}^{N} p_i \log_2 p_i \,,$$

gdzie przyjmujemy $0 \cdot \ln 0 \;=\; 0$, a k jest stałą Boltzmanna.

Entropia statystyczna

- Niech $P = (p_1, \ldots, p_N)$ będzie prawdopodobieństwem (N może być nieskończone!)

Definicja

Entropię definiuje się jako

$$S(P) \;=\; -k\sum_{i=1}^{N} p_i \ln p_i \quad \text{albo} \quad S(P) \;=\; -\sum_{i=1}^{N} p_i \log_2 p_i \,,$$

gdzie przyjmujemy $0 \cdot \ln 0 \;=\; 0$, a k jest stałą Boltzmanna.

- $N = 2$

$$S(p) = -k[p \ln p + (1-p)\ln(1-p)]\,, \qquad 0 \leqslant p \leqslant 1$$

Entropia statystyczna

- Niech $P = (p_1, \ldots, p_N)$ będzie prawdopodobieństwem (N może być nieskończone!)

Definicja

Entropię definiuje się jako

$$S(P) = -k \sum_{i=1}^{N} p_i \ln p_i \quad \text{albo} \quad S(P) = -\sum_{i=1}^{N} p_i \log_2 p_i \,,$$

gdzie przyjmujemy $0 \cdot \ln 0 = 0$, a k jest stałą Boltzmanna.

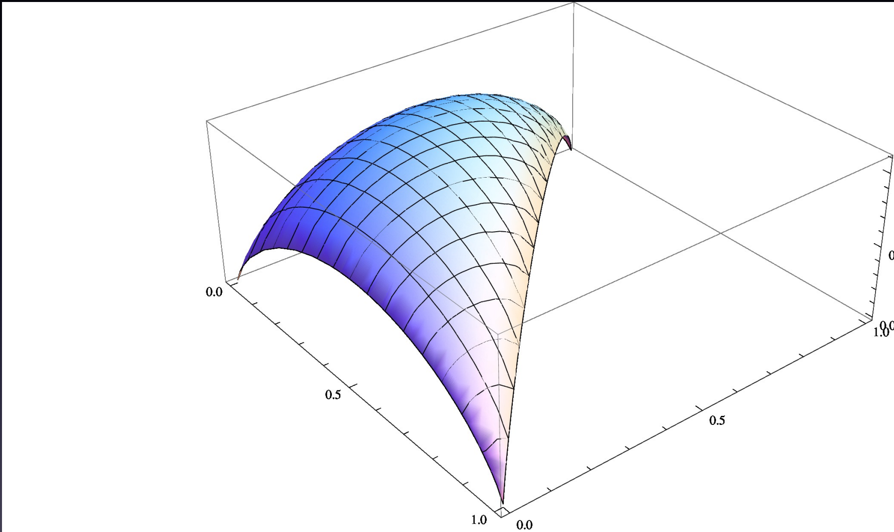
- $N = 2$

$$S(p) = -k[p \ln p + (1-p)\ln(1-p)]\,, \qquad 0 \leqslant p \leqslant 1$$

- $N = 3$

$$S(p_1, p_2) = -k[p_1 \ln p_1 + p_2 \ln p_2 + (1-p_1-p_2)\ln(1-p_1-p_2)]\,, \quad 0 \leqslant p_1, p_2 \leqslant 1$$

Wykresy entropii



Własności entropii

Własności entropii

- Własności entropii:

- Własności entropii:
  1. $0 \leqslant S(P) \leqslant k \ln N$

Własności entropii

- Własności entropii:
  1. $0 \leqslant S(P) \leqslant k \ln N$

     $S(P) = 0 \iff P = (0, \ldots, 1, \ldots, 0)$

     $S(P) = k \ln N \iff P = (\frac{1}{N}, \ldots, \frac{1}{N})$

     (dowód w oparciu o tw. Lagrange'a)

Tw. Lagrange'a

Tw. Lagrange'a

Tw. Lagrange'a

Jeśli funkcja $f(x_1,\ldots,x_n)$ ma w pewnym punkcie ekstremum warunkowe przy warunkach $W_i(x_1,\ldots,x_n)=0$, $i=1,\ldots,r<n$, to funkcja

$$g(x_1,\ldots,x_n) \;=\; f(x_1,\ldots,x_n)-\sum_{i=1}^{r}\lambda_i W_i(x_1,\ldots,x_n)$$

ma w tym punkcie ekstremum lokalne przy pewnych wartościach mnożników Lagrange'a λ_i.

Sympleks

Sympleks

- Zbiór wszystkich rozkładów prawdopodobieństw

$$\mathcal{P}_N = \left\{ (p_1, \ldots, p_N) \, ; \, 0 \leqslant p_i \leqslant 1 \, , \sum_{i=1}^{N} p_i = 1 \right\}$$ tworzy sympleks

- Sympleks $\mathcal{P}_N$ jest zbiorem wypukłym, tzn.

$$\forall \; P, Q \in \mathcal{P}_N \quad \forall \; \alpha \in [0,1] \quad \alpha P + (1-\alpha) Q \; \in \; \mathcal{P}_N$$

Sympleks

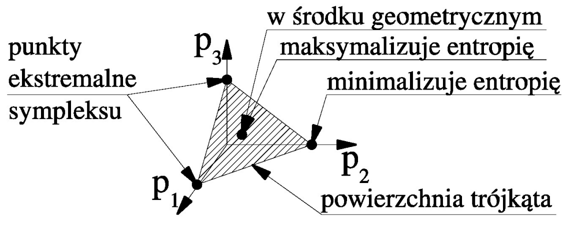
- Zbiór wszystkich rozkładów prawdopodobieństw

$$\mathcal{P}_N = \left\{ (p_1, \ldots, p_N) \, ; \, 0 \leqslant p_i \leqslant 1 \, , \sum_{i=1}^{N} p_i = 1 \right\}$$ tworzy sympleks

- Sympleks $\mathcal{P}_N$ jest zbiorem wypukłym, tzn.

$$\forall \; P, Q \in \mathcal{P}_N \quad \forall \; \alpha \in [0,1] \quad \alpha P + (1-\alpha) Q \; \in \; \mathcal{P}_N$$

Sympleks przy $N = 3$



Własności entropii

Własności entropii

3. Entropia $S(P)$ jest funkcją wklęsłą (wypukłą w górę) na dowolnym wypukłym podzbiorze $\mathcal{P} \subset \mathcal{P}_N$, tzn.

$$\forall\, \alpha \in [0,1] \quad S(\alpha P + (1-\alpha)Q) \;\geqslant\; \alpha S(P) + (1-\alpha) S(Q)\,.$$

Równość zachodzi wtedy i tylko wtedy gdy $P = Q$.

Własności entropii

3. Entropia $S(P)$ jest funkcją wklęsłą (wypukłą w górę) na dowolnym wypukłym podzbiorze $\mathcal{P} \subset \mathcal{P}_N$, tzn.

$$\forall\, \alpha \in [0,1] \quad S(\alpha P + (1-\alpha)Q) \;\geqslant\; \alpha S(P) + (1-\alpha) S(Q)\,.$$

   Równość zachodzi wtedy i tylko wtedy gdy $P = Q$.

4. Jeśli na zbiorze wypukłym $\mathcal{P} \subset \mathcal{P}_N$ entropia jest ograniczona z góry, to osiąga maksimum na dokładnie jednym rozkładzie prawdopodobieństwa $P^* = (p_1^*, \ldots, p_N^*)$,

$$S(P^*) \;=\; \sup_{P \in \mathcal{P}} S(P)\,.$$

Własności entropii

3. Entropia $S(P)$ jest funkcją wklęsłą (wypukłą w górę) na dowolnym wypukłym podzbiorze $\mathcal{P} \subset \mathcal{P}_N$, tzn.

$$\forall\, \alpha \in [0,1] \quad S(\alpha P + (1-\alpha)Q) \;\geqslant\; \alpha S(P) + (1-\alpha) S(Q)\,.$$

   Równość zachodzi wtedy i tylko wtedy gdy $P = Q$.

4. Jeśli na zbiorze wypukłym $\mathcal{P} \subset \mathcal{P}_N$ entropia jest ograniczona z góry, to osiąga maksimum na dokładnie jednym rozkładzie prawdopodobieństwa $P^* = (p_1^*, \ldots, p_N^*)$,

$$S(P^*) \;=\; \sup_{P \in \mathcal{P}} S(P)\,.$$

Definicja

Rozkładem reprezentatywnym P^* zbioru $\mathcal{P}$ nazywamy jedyny rozkład prawdopodobieństwa maksymalizujący entropię.

Makrostan ze względu na zmienną losową

Makrostan ze względu na zmienną losową

Makrostan

Makrostanem ze względu na dyskretną zmienną losową f nazywamy następujący wypukły podzbiór sympleksu $\mathcal{P}_N$

$$\mathcal{P}_N \supset \mathcal{K}_f = \left\{ P = (p_1, \ldots, p_N)\,;\, p_i \geqslant 0\,,\, \sum_{i=1}^{N} p_i = 1\,,\, \langle f \rangle = m \right\},$$

Makrostan ze względu na zmienną losową

Makrostan

Makrostanem ze względu na dyskretną zmienną losową f nazywamy następujący wypukły podzbiór sympleksu $\mathcal{P}_N$

$$\mathcal{P}_N \supset \mathcal{K}_f = \left\{ P = (p_1, \ldots, p_N)\,;\, p_i \geqslant 0\,,\, \sum_{i=1}^{N} p_i = 1\,,\, \langle f \rangle = m \right\},$$

Tw. o postaci rozkładu reprezentatywnego dla makrostanu

Rozkład reprezentatywny $P^* = (p_1^*, \ldots, p_N^*)$ makrostanu $\mathcal{K}_f$ ma postać:

$$p_i^* = Z^{-1}(\lambda) e^{-\lambda f_i}\,,$$

gdzie funkcję $Z(\lambda) = \sum_{i=1}^{N} e^{-\lambda f_i}$ nazywamy sumą statystyczną oraz równość

$$-\frac{\partial \ln Z(\lambda)}{\partial \lambda} = m$$

jednoznacznie określa funkcję $\lambda(m)$.

Makrostan ze względu na kilka zmiennych losowych

Makrostan ze względu na kilka zmiennych losowych

Przykład: Znaleźć rozkład reprezentatywny dla makrostanu

$$K_f \;=\; \left\{ P = (p_1, p_2, \ldots); p_i \geqslant 0\,,\; \sum_{i=1}^{\infty} p_i = 1\,,\; \langle f \rangle = m \right\}$$

jeśli wartościami zmiennej losowej f są kolejne liczby naturalne, tzn. $f_i = i$, $i = 1, 2, \ldots$.

Definicja

Makrostanem ze względu na dyskretne zmienne losowe $f^1, \ldots, f^r$ nazywamy następujący wypukły podzbiór sympleksu $\mathcal{P}_N$

$$\mathcal{P}_N \supset \mathcal{K}_{f^1,\ldots,f^r} \;=\; \Big\{ (p_1, \ldots, p_N)\,;\; p_i \geqslant 0\,,\; \sum_{i=1}^{N} p_i = 1\,,\; \langle f^j \rangle = m_j\,, \;\; j = 1, \ldots, r \Big\}\,, \qquad f_i^j \;:=\; f^j(\omega_i)\,,$$

Twierdzenie o rozkładzie reprezentatywnym

Twierdzenie

Rozkład reprezentatywny $P^* = (p_1^*, \ldots, p_N^*)$ makrostanu $\mathcal{K}_{f^1,\ldots,f^r}$ ma postać:

$$p_i^* = Z^{-1}(\lambda_1, \ldots, \lambda_r) \exp\left[-\sum_{j=1}^{r} \lambda_j f_i^j\right],$$

gdzie funkcję $Z(\lambda_1, \ldots, \lambda_r) = \sum_{i=1}^{N} \exp\left[-\sum_{j=1}^{r} \lambda_j f_i^j\right]$ nazywamy sumą statystyczną

oraz

$$-\frac{\partial \ln Z(\lambda_1, \ldots, \lambda_r)}{\partial \lambda_j} = m_j\,, \quad j = 1, \ldots, r\,.$$

Znaczenie entropii

Znaczenie entropii

- Informacja — zajście zdarzenia ω_k dostarcza informacji

$$I_k(P) = \log_2 \frac{1}{p(\omega_k)} = -\log_2 p_k$$

Znaczenie entropii

- Informacja — zajście zdarzenia ω_k dostarcza informacji

$$I_k(P) = \log_2 \frac{1}{p(\omega_k)} = -\log_2 p_k$$

- Zdarzenia mniej prawdopodobne dostarczają więcej informacji!

Znaczenie entropii

- Informacja — zajście zdarzenia ω_k dostarcza informacji

$$I_k(P) = \log_2 \frac{1}{p(\omega_k)} = -\log_2 p_k$$

- Zdarzenia mniej prawdopodobne dostarczają więcej informacji!
- Entropia określa średnią informację jaka może być uzyskana w wyniku jakiegoś doświadczenia losowego

$$S(P) = \langle I(P) \rangle = -\sum_k p_k \log_2 p_k$$

Znaczenie entropii

- Informacja — zajście zdarzenia ω_k dostarcza informacji

$$I_k(P) = \log_2 \frac{1}{p(\omega_k)} = -\log_2 p_k$$

- Zdarzenia mniej prawdopodobne dostarczają więcej informacji!
- Entropia określa średnią informację jaka może być uzyskana w wyniku jakiegoś doświadczenia losowego

$$S(P) = \langle I(P) \rangle = -\sum_k p_k \log_2 p_k$$

- Jeśli P jest rozkładem zmiennej losowej f, to $S(P)$ określa (średnią) ilość informacji zawartą w f.

Inne rodzaje entropii

Inne rodzaje entropii

- Miara podobieństwa dwóch rozkładów P i Q (Q absolutnie ciągły względem P) — entropia wzajemna

$$S(P||Q) = \sum_k p_k \log_2 \frac{p_k}{q_k} = -S(P) - \sum_k p_k \log_2 q_k$$

Inne rodzaje entropii

- Miara podobieństwa dwóch rozkładów P i Q (Q absolutnie ciągły względem P) — entropia wzajemna

$$S(P||Q) = \sum_k p_k \log_2 \frac{p_k}{q_k} = -S(P) - \sum_k p_k \log_2 q_k$$

- Entropia wzajemna jest nieujemna, przy czym wartość 0 przyjmuje tylko dla identycznych rozkładów $P = Q$

$$S(P||Q) \geqslant 0$$

Inne rodzaje entropii

- Miara podobieństwa dwóch rozkładów P i Q (Q absolutnie ciągły względem P) — entropia wzajemna

$$S(P||Q) = \sum_k p_k \log_2 \frac{p_k}{q_k} = -S(P) - \sum_k p_k \log_2 q_k$$

- Entropia wzajemna jest nieujemna, przy czym wartość 0 przyjmuje tylko dla identycznych rozkładów $P = Q$

$$S(P||Q) \geqslant 0$$

- Entropia rozkładu łącznego $u_{k\ell} = (P \times Q)(\omega_k, \eta_\ell)$

$$S(P \times Q) = -\sum_{k,\ell} u_{k\ell} \log_2 u_{k\ell}$$

Inne rodzaje entropii

- Miara podobieństwa dwóch rozkładów P i Q (Q absolutnie ciągły względem P) — entropia wzajemna

$$S(P||Q) = \sum_k p_k \log_2 \frac{p_k}{q_k} = -S(P) - \sum_k p_k \log_2 q_k$$

- Entropia wzajemna jest nieujemna, przy czym wartość 0 przyjmuje tylko dla identycznych rozkładów $P = Q$

$$S(P||Q) \geqslant 0$$

- Entropia rozkładu łącznego $u_{k\ell} = (P \times Q)(\omega_k, \eta_\ell)$

$$S(P \times Q) = -\sum_{k,\ell} u_{k\ell} \log_2 u_{k\ell}$$

- Entropia warunkowa — ilość informacji o $P \times Q$ jaka może być uzyskana na podstawie znajomości rozkładu Q

$$S(P|Q) = S(P \times Q) - S(Q)$$

Entropia wzajemna

Entropia wzajemna

- Wspólna informacja zawarta w P i Q

$$S(P:Q) = S(P) + S(Q) - S(P \times Q)$$

Entropia wzajemna

- Wspólna informacja zawarta w P i Q

$$S(P:Q) = S(P) + S(Q) - S(P \times Q)$$

- Własności

$$S(P:Q) = S(P) - S(P|Q)$$

Entropia wzajemna

- Wspólna informacja zawarta w P i Q

$$S(P:Q) = S(P) + S(Q) - S(P \times Q)$$

- Własności

$$S(P:Q) = S(P) - S(P|Q)$$

$$S(P:Q) = S(P \times Q || PQ)$$

Entropia w przypadku ciągłej zmiennej losowej

Entropia w przypadku ciągłej zmiennej losowej

- Ciągłej zmiennej losowej f określonej na podzbiorach $\Omega \subset \mathbb{R}^n$ odpowiada gęstość rozkładu $\rho(x_1, \ldots, x_n) \equiv \rho(\underline{x})$

Entropia w przypadku ciągłej zmiennej losowej

- Ciągłej zmiennej losowej f określonej na podzbiorach $\Omega \subset \mathbb{R}^n$ odpowiada gęstość rozkładu $\rho(x_1, \ldots, x_n) \equiv \rho(\underline{x})$

Entropia

Entropię definiujemy w tym przypadku jako

$$S(\rho) \; = \; -k \int_{\Omega} \rho(\underline{x}) \ln \rho(\underline{x}) d\underline{x} \, ,$$

Entropia w przypadku ciągłej zmiennej losowej

- Ciągłej zmiennej losowej f określonej na podzbiorach $\Omega \subset \mathbb{R}^n$ odpowiada gęstość rozkładu $\rho(x_1, \ldots, x_n) \equiv \rho(\underline{x})$

Entropia

Entropię definiujemy w tym przypadku jako

$$S(\rho) \; = \; -k \int_\Omega \rho(\underline{x}) \ln \rho(\underline{x}) d\underline{x} \,,$$

Makrostan

Makrostanem ze względu na ciągłe zmienne losowe $f^1, \ldots, f^r$ nazywamy następujący zbiór:

$$\mathcal{K}_{f^1,\ldots,f^r} = \left\{ \rho : \Omega \to \mathbb{R}^+ ; \; \int_\Omega \rho(\underline{x}) d\underline{x} = 1 \,, \; E(f^j) = m_j \,, \quad j = 1, \ldots, r \right\},$$

Twierdzenie o rozkładzie reprezentatywnym

Twierdzenie

Gęstość rozkładu reprezentatywnego $\rho^*(\underline{x})$ makrostanu $\mathcal{K}_{f^1,\ldots,f^r}$ ma postać:

$$\rho^*(\underline{x}) \;=\; Z^{-1}(\lambda_1,\ldots,\lambda_r)\exp\left[-\sum_{j=1}^{r}\lambda_j f^j(\underline{x})\right],$$

gdzie suma statystyczna ma postać

$$Z(\lambda_1,\ldots,\lambda_r) \;=\; \int_\Omega \exp\left[-\sum_{j=1}^{r}\lambda_j f^j(\underline{x})\right] d\underline{x}$$

oraz

$$-\frac{\partial \ln Z(\lambda_1,\ldots,\lambda_r)}{\partial\lambda_j} \;=\; m_j\,,\quad j=1,\ldots,r\,.$$

gdzie $m_j \;=\; \int_\Omega f^j(\underline{x})\rho(\underline{x})d\underline{x}$.