Enrolamento e quantização de distribuições de probabilidade por reticulados

Fábio C. C. Meneghetti

IMECC — Unicamp

7 de junho de 2022

Introdução •00000

• seja $\Lambda \subset \mathbb{R}^n$ reticulado de posto completo $(\dim \Lambda = n)$



Introdução •00000

- seja $\Lambda \subset \mathbb{R}^n$ reticulado de posto completo $(\dim \Lambda = n)$
- ullet um domínio fundamental é um conjunto mensurável ${\mathcal D}$ tal que

Introdução •00000

- seja $\Lambda \subset \mathbb{R}^n$ reticulado de posto completo $(\dim \Lambda = n)$
- ullet um domínio fundamental é um conjunto mensurável ${\cal D}$ tal que



- seja $\Lambda \subset \mathbb{R}^n$ reticulado de posto completo (dim $\Lambda = n$)
- ullet um domínio fundamental é um conjunto mensurável ${\cal D}$ tal que

- seja $\Lambda \subset \mathbb{R}^n$ reticulado de posto completo (dim $\Lambda = n$)
- ullet um domínio fundamental é um conjunto mensurável ${\cal D}$ tal que

• Exemplos:

Introdução •00000

- seja $\Lambda \subset \mathbb{R}^n$ reticulado de posto completo $(\dim \Lambda = n)$
- ullet um domínio fundamental é um conjunto mensurável ${\cal D}$ tal que

 - $(\lambda + \mathcal{D}) \cap (\lambda' + \mathcal{D}) = \emptyset, \quad \lambda \neq \lambda' \in \Lambda$
- Exemplos:

Introdução

• paralelotopo fundamental com relação a uma base $\beta = \{b_1, \dots, b_n\}$:

$$\mathcal{P}(\beta) = \{x_1b_1 + \dots x_nb_n : x_i \in [0,1)\}$$

- seja $\Lambda \subset \mathbb{R}^n$ reticulado de posto completo (dim $\Lambda = n$)
- ullet um domínio fundamental é um conjunto mensurável ${\cal D}$ tal que

 - $(\lambda + \mathcal{D}) \cap (\lambda' + \mathcal{D}) = \emptyset, \quad \lambda \neq \lambda' \in \Lambda$
- Exemplos:

Introdução 000000

• paralelotopo fundamental com relação a uma base $\beta = \{b_1, \dots, b_n\}$:

$$\mathcal{P}(\beta) = \left\{ x_1 b_1 + \dots x_n b_n : x_i \in [0, 1) \right\}$$

 região de Voronói, pode ser reduzida a um domínio fundamental cortando algumas bordas:

$$\mathcal{V}(\Lambda) = \left\{ x \in \mathbb{R}^n : ||x|| \le ||x + \lambda||, \quad \forall \lambda \in \Lambda \right\}$$

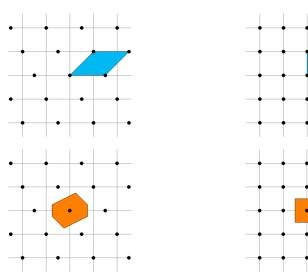


Figura: Paralelotopo fundamental e região de Voronói, respectivamente.



Introdução ○○●○○○

• o quociente de \mathbb{R}^n por Λ é

$$\mathbb{R}^n/\Lambda := \{x + \Lambda : x \in \mathbb{R}^n\},$$

que é munido de uma função projeção $\pi \colon \mathbb{R}^n \to \mathbb{R}^n/\Lambda$,

$$\pi(x) = x + \Lambda.$$

Proposição

Introdução ○○○●○○

O mapa restrito $\pi|_{\mathcal{D}} \colon \mathcal{D} \to \mathbb{R}^n/\Lambda$ é bijeção, para qualquer região fundamental \mathcal{D} .



Proposição

Introdução ○○○●○○

O mapa restrito $\pi|_{\mathcal{D}} \colon \mathcal{D} \to \mathbb{R}^n/\Lambda$ é bijeção, para qualquer região fundamental \mathcal{D} .



Proposição

O mapa restrito $\pi|_{\mathcal{D}} \colon \mathcal{D} \to \mathbb{R}^n/\Lambda$ é bijeção, para qualquer região fundamental \mathcal{D} .

Demonstração.



Proposição

O mapa restrito $\pi|_{\mathcal{D}} \colon \mathcal{D} \to \mathbb{R}^n/\Lambda$ é bijeção, para qualquer região fundamental \mathcal{D} .

Demonstração.

1 Injetividade: $\pi|_{\mathcal{D}}(x) = \pi|_{\mathcal{D}}(y) \iff x + \Lambda = y + \Lambda \iff x - y \in \Lambda \iff x + 0 = y + \lambda \text{ para algum } \lambda \in \Lambda.$ Ter $\lambda \neq 0$ iria contradizer $(\mathcal{D} + 0) \cap (\mathcal{D} + \lambda) = \emptyset$, portanto x = y.

Proposição

O mapa restrito $\pi|_{\mathcal{D}} \colon \mathcal{D} \to \mathbb{R}^n/\Lambda$ é bijeção, para qualquer região fundamental D.

Demonstração.

- **1** Injetividade: $\pi|_{\mathcal{D}}(x) = \pi|_{\mathcal{D}}(y) \iff x + \Lambda = y + \Lambda \iff x y \in \mathcal{D}(x)$ $\Lambda \iff x+0=y+\lambda$ para algum $\lambda \in \Lambda$. Ter $\lambda \neq 0$ iria contradizer $(\mathcal{D}+0)\cap(\mathcal{D}+\lambda)=\emptyset$, portanto x=y.
- **Sobrejetividade:** tome $(x + \Lambda) \in \mathbb{R}^n / \Lambda$. De $\mathbb{R}^n = \bigcup_{\lambda \in \Lambda} (\lambda + \mathcal{D})$, existem $\bar{x} \in \mathcal{D}$, $\lambda \in \Lambda$ tais que $x = \bar{x} + \lambda$. Assim

$$(x + \Lambda) = (\bar{x} + \Lambda) = \pi|_{\mathcal{D}}(\bar{x}).$$

Introdução 000000

em outras palavras, cada classe $x + \Lambda$ possui um único representante em \mathcal{D} .

- em outras palavras, cada classe $x + \Lambda$ possui um único representante em \mathcal{D} .
- essa identificação é mensurável: não altera o volume de subconjuntos, e esse volume é o mesmo para qualquer domínio fundamental escolhido.

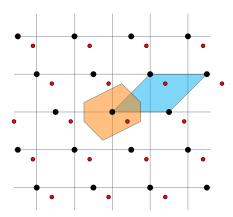


Figura: Classe lateral $x+\Lambda$ interseccionando cada região fundamental em apenas um ponto.

Introdução 00000●

Distribuição enrolada

• seja $p \colon \mathbb{R}^n \to \mathbb{R}_+$ função tal que $\int_{\mathbb{R}^n} p(x) \, \mathrm{d}x = 1$ (chamada função densidade de probabilidade)



Distribuição enrolada

- seja $p: \mathbb{R}^n \to \mathbb{R}_+$ função tal que $\int_{\mathbb{R}^n} p(x) \, \mathrm{d}x = 1$ (chamada função densidade de probabilidade)
- a distribuição enrolada por um reticulado Λ é definida como $p_{\pi} \colon \mathcal{D} \to \mathbb{R}_{+}$

$$p_{\pi}(x) = \sum_{\lambda \in \Lambda} p(x + \lambda)$$

- seja $p \colon \mathbb{R}^n \to \mathbb{R}_+$ função tal que $\int_{\mathbb{R}^n} p(x) \, \mathrm{d}x = 1$ (chamada função densidade de probabilidade)
- a distribuição enrolada por um reticulado Λ é definida como $p_{\pi} : \mathcal{D} \to \mathbb{R}_{+}$,

$$p_{\pi}(x) = \sum_{\lambda \in \Lambda} p(x + \lambda)$$

• em outras palavras, somamos a probabilidade para cada classe $(x + \Lambda) \in \mathbb{R}^n/\Lambda$

- seja $p \colon \mathbb{R}^n \to \mathbb{R}_+$ função tal que $\int_{\mathbb{R}^n} p(x) \, \mathrm{d}x = 1$ (chamada função densidade de probabilidade)
- a distribuição enrolada por um reticulado Λ é definida como $p_{\pi} : \mathcal{D} \to \mathbb{R}_{+}$,

$$p_{\pi}(x) = \sum_{\lambda \in \Lambda} p(x + \lambda)$$

- em outras palavras, somamos a probabilidade para cada classe $(x + \Lambda) \in \mathbb{R}^n/\Lambda$
- ullet alternativamente, poderia ser vista como distribuição sobre \mathbb{R}^n/Λ

① Normal univariada. $p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right), \quad \Lambda = k\mathbb{Z}$

$$p_{\pi}(x) = \frac{1}{\sqrt{2\pi}\sigma} \sum_{z \in \mathbb{Z}} \exp\left(\frac{-(x + kz - \mu)^2}{2\sigma^2}\right)$$

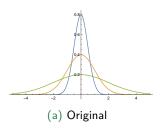
① Normal univariada. $p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right), \quad \Lambda = k\mathbb{Z}$

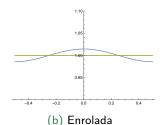
$$p_{\pi}(x) = \frac{1}{\sqrt{2\pi}\sigma} \sum_{z \in \mathbb{Z}} \exp\left(\frac{-(x + kz - \mu)^2}{2\sigma^2}\right)$$

Exemplos

1 Normal univariada. $p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(\frac{-(x-\mu)^2}{2\sigma^2}\right), \quad \Lambda = k\mathbb{Z}$

$$ho_\pi(x) = rac{1}{\sqrt{2\pi}\sigma} \sum_{z \in \mathbb{Z}} \exp\left(rac{-(x+kz-\mu)^2}{2\sigma^2}
ight)$$





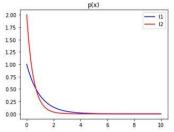
2 Exponencial. $p(x) = \nu e^{-\nu x}, x > 0, \Lambda = k\mathbb{Z}$

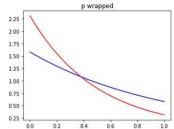
$$p_{\pi}(y) = \frac{\nu e^{-\nu y}}{1 - e^{-k\nu}}$$

2 Exponencial. $p(x) = \nu e^{-\nu x}, x > 0, \Lambda = k\mathbb{Z}$

$$p_{\pi}(y) = \frac{\nu e^{-\nu y}}{1 - e^{-k\nu}}$$

$$p_{\pi}(y) = \frac{\nu e^{-\nu y}}{1 - e^{-k\nu}}$$





Aplicações

Enrolamento 000€0

Estatística direcional. estuda a estatística de dados que são circulares, definidos no círculo, esfera (\mathbb{S}^n) ou toro (\mathbb{T}^n) — esta último é nosso caso.

Exemplos:



Aplicações

Enrolamento

Estatística direcional. estuda a estatística de dados que são circulares, definidos no círculo, esfera (\mathbb{S}^n) ou toro (\mathbb{T}^n) — esta último é nosso caso.

Exemplos:

• a frequência de um evento em função do dia do ano é um dado circular

Enrolamento 00000

Aplicações

Estatística direcional. estuda a estatística de dados que são circulares, definidos no círculo, esfera (\mathbb{S}^n) ou toro (\mathbb{T}^n) — esta último é nosso caso.

Exemplos:

- a frequência de um evento em função do dia do ano é um dado circular
- a distribuição de substâncias na superfície de um planeta é um dado esférico

Aplicações

Estatística direcional. estuda a estatística de dados que são circulares, definidos no círculo, esfera (\mathbb{S}^n) ou toro (\mathbb{T}^n) — esta último é nosso caso.

Exemplos:

- a frequência de um evento em função do dia do ano é um dado circular
- a distribuição de substâncias na superfície de um planeta é um dado esférico
- a distribuição conjunta de quaisquer dois dados circulares é um dado no toro $\mathbb{T}^2=\mathbb{S}^1\times\mathbb{S}^1$



Codificação em canais AWGN e Wiretap.

2 Codificação em canais AWGN e Wiretap.

• O fator de achamento é definido como a distância L^{∞} entre uma distribuição enrolada e uma uniforme em \mathcal{D} :

$$\epsilon_{\Lambda}(p) = \sup_{x \in \mathcal{D}} \left| p_w(x) - \frac{1}{\det \Lambda} \right|$$

• O fator de achamento é definido como a distância L^{∞} entre uma distribuição enrolada e uma uniforme em \mathcal{D} :

$$\epsilon_{\Lambda}(p) = \sup_{x \in \mathcal{D}} \left| p_w(x) - \frac{1}{\det \Lambda} \right|$$

 quando temos um canal AWGN ou Wiretap com ruído gaussiano, o fator de achatamento pode ser usado construír códigos com eficiência máxima (atingem a capacidade de Shannon)

- Codificação em canais AWGN e Wiretap.
 - O fator de achamento é definido como a distância L^{∞} entre uma distribuição enrolada e uma uniforme em \mathcal{D} :

$$\epsilon_{\Lambda}(p) = \sup_{x \in \mathcal{D}} \left| p_w(x) - \frac{1}{\det \Lambda} \right|$$

- quando temos um canal AWGN ou Wiretap com ruído gaussiano, o fator de achatamento pode ser usado construír códigos com eficiência máxima (atingem a capacidade de Shannon)
- Criptografia baseada em reticulados.

Codificação em canais AWGN e Wiretap.

• O fator de achamento é definido como a distância L^{∞} entre uma distribuição enrolada e uma uniforme em \mathcal{D} :

$$\epsilon_{\Lambda}(p) = \sup_{x \in \mathcal{D}} \left| p_w(x) - \frac{1}{\det \Lambda} \right|$$

 quando temos um canal AWGN ou Wiretap com ruído gaussiano, o fator de achatamento pode ser usado construír códigos com eficiência máxima (atingem a capacidade de Shannon)

Criptografia baseada em reticulados.

• O parâmetro de suavização, equivalente ao fator de achatamento, é um parâmetro de garantia de seguranca e confiabilidade na criptografia pós-quântica baseada em reticulados.

Distribuições quantizadas

• pode ser considerada uma operação dual ao enrolamento.



- pode ser considerada uma operação dual ao enrolamento.
- a distribuição quantizada por um reticulado Λ e uma região fundamental \mathcal{D} é a distribuição $p_{\mathcal{Q}}: \Lambda \to \mathbb{R}_+$,

$$p_{\mathcal{Q}}(\lambda) = \int_{\mathcal{D}} p(x + \lambda) \, dx.$$

- pode ser considerada uma operação dual ao enrolamento.
- a distribuição quantizada por um reticulado Λ e uma região fundamental \mathcal{D} é a distribuição $p_{\mathcal{Q}} \colon \Lambda \to \mathbb{R}_+$,

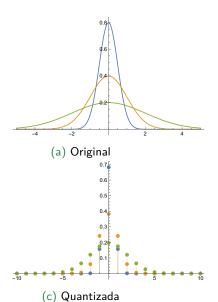
$$p_{\mathcal{Q}}(\lambda) = \int_{\mathcal{D}} p(x + \lambda) dx.$$

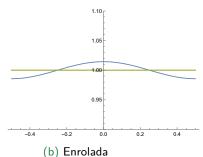
• se o enrolamento é somar $p(x + \lambda)$ em λ , a quantização é "somar" (integrar) $p(x + \lambda)$ em x.

- pode ser considerada uma operação dual ao enrolamento.
- a distribuição quantizada por um reticulado Λ e uma região fundamental \mathcal{D} é a distribuição $p_{\mathcal{Q}}: \Lambda \to \mathbb{R}_+$,

$$p_{\mathcal{Q}}(\lambda) = \int_{\mathcal{D}} p(x + \lambda) dx.$$

- se o enrolamento é somar $p(x + \lambda)$ em λ , a quantização é "somar" (integrar) $p(x + \lambda)$ em x.
- é uma forma de transformar uma distribuição de probabilidade contínua em uma discreta.







Descrição em variáveis aleatórias

• uma função densidade p(x) descreve a distribuição de uma variável aleatória X, com

$$\mathbb{P}[X \in A] = \int_A p(x) \, \mathrm{d}x$$

Notação: $X \sim p$

• uma função densidade p(x) descreve a distribuição de uma variável aleatória X, com

$$\mathbb{P}[X \in A] = \int_A p(x) \, \mathrm{d}x$$

Notação: $X \sim p$

• temos duas funções: o enrolamento $\pi: \mathbb{R}^n \to \mathcal{D}$ e a quantização $\mathcal{Q} \colon \mathbb{R}^n \to \Lambda$, que são definidas por

$$\pi(y + \lambda) = y,$$
 $Q(y + \lambda) = \lambda,$

para todo $v \in \mathcal{D}$. $\lambda \in \Lambda$



$$X \sim p \implies egin{cases} X_{\pi} \coloneqq \pi(X) \sim p_{\pi} \ X_{\mathcal{Q}} \coloneqq \mathcal{Q}(X) \sim p_{\mathcal{Q}} \end{cases}$$

$$X \sim p \implies \begin{cases} X_{\pi} := \pi(X) \sim p_{\pi} \\ X_{\mathcal{Q}} := \mathcal{Q}(X) \sim p_{\mathcal{Q}} \end{cases}$$

• se definimos X_{π} e $X_{\mathcal{O}}$ dessa forma, temos que $X=X_{\pi}+X_{\mathcal{O}}$

$$X \sim p \implies \begin{cases} X_{\pi} := \pi(X) \sim p_{\pi} \\ X_{\mathcal{Q}} := \mathcal{Q}(X) \sim p_{\mathcal{Q}} \end{cases}$$

- se definimos X_{π} e $X_{\mathcal{Q}}$ dessa forma, temos que $X=X_{\pi}+X_{\mathcal{Q}}$
- essa descrição estabelece uma relação entre as distribuições enrolada e quantizada, e nos permite investigar propriedades dessa relação

Esperança

• a esperança (média) de uma variável aleatória X (ou de uma distribuição p) é $E[X] := \int_{\mathbb{R}^n} x p(x) dx$



Esperança

- a esperança (média) de uma variável aleatória X (ou de uma distribuição p) é $E[X] := \int_{\mathbb{R}^n} x p(x) \, dx$
- as esperanças enrolada e quantizada são $E[X_{\pi}] = \int_{\mathcal{D}} x p_{\pi}(x) \, \mathrm{d}x$ e $E[X_{\mathcal{O}}] = \sum_{\lambda \in \Lambda} \lambda p_{\mathcal{O}}(\lambda)$



- a esperança (média) de uma variável aleatória X (ou de uma distribuição p) é $E[X] := \int_{\mathbb{R}^n} x p(x) dx$
- as esperanças enrolada e quantizada são $E[X_{\pi}] = \int_{\mathcal{D}} x p_{\pi}(x) dx$ e $E[X_{\mathcal{O}}] = \sum_{\lambda \in \Lambda} \lambda p_{\mathcal{O}}(\lambda)$
- disso, temos que $E[X] = E[X_{\pi}] + E[X_{\mathcal{O}}]$

$$H[X] = -\int_{\mathbb{R}^n} p(x) \log p(x) dx,$$
 $H[X_{\pi}] = -\int_{\mathcal{D}} p_{\pi}(y) \log p_{\pi}(y) dy,$
 $H[X_{\mathcal{Q}}] = -\sum_{\lambda \in \Lambda} p_{\mathcal{Q}}(\lambda) \log p_{\mathcal{Q}}(\lambda)$

• temos $H[X] < H[X_{\pi}] + H[X_{\mathcal{O}}]$

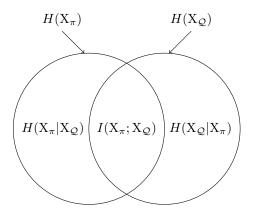


$$H[X] = -\int_{\mathbb{R}^n} p(x) \log p(x) dx,$$
 $H[X_{\pi}] = -\int_{\mathcal{D}} p_{\pi}(y) \log p_{\pi}(y) dy,$
 $H[X_{\mathcal{Q}}] = -\sum_{\lambda \in \Lambda} p_{\mathcal{Q}}(\lambda) \log p_{\mathcal{Q}}(\lambda)$

- temos $H[X] < H[X_{\pi}] + H[X_{\mathcal{O}}]$
- mais precisamente, $H[X] = H[X_{\pi}] + H[X_{\mathcal{O}}] I(X_{\pi}; X_{\mathcal{O}})$

Propriedades informacionais

000000000



Informação mútua

 estamos interessados em entender como essa informação mútua se comporta



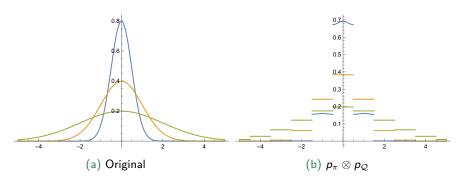
Informação mútua

- estamos interessados em entender como essa informação mútua se comporta
- ela pode ser caracterizada através da divergência de Kullback-Leibler como

$$I(X_{\pi}; X_{\mathcal{Q}}) = D_{\mathsf{KL}}(p||p_{\pi} \otimes p_{\mathcal{Q}}),$$

onde p é a distribuição original e $p_\pi \otimes p_\mathcal{Q}(x) = p_\pi(\pi(x))p_\mathcal{Q}(\mathcal{Q}(x))$

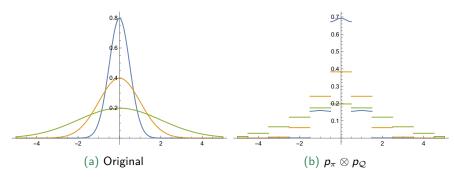
Distribuições normais com média 0



• σ grande $\implies p_{\pi} pprox rac{1}{\det \Lambda}$ e $p_{\mathcal{Q}} pprox \det \Lambda \cdot p \implies p_{\pi} \otimes p_{\mathcal{Q}}$ aproxima p

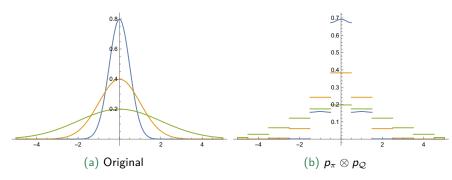


Distribuições normais com média 0



- σ grande $\implies p_{\pi} \approx \frac{1}{\det \Lambda}$ e $p_{\mathcal{Q}} \approx \det \Lambda \cdot p \implies p_{\pi} \otimes p_{\mathcal{Q}}$ aproxima p
- σ pequeno $\implies p_{\mathcal{Q}} \approx \delta_0$ e $p_{\pi} \approx p|_{\mathcal{D}} \implies p_{\pi} \otimes p_{\mathcal{Q}}$ aproxima p

Distribuições normais com média 0



- σ grande $\implies p_{\pi} pprox rac{1}{\det \Lambda}$ e $p_{\mathcal{Q}} pprox \det \Lambda \cdot p \implies p_{\pi} \otimes p_{\mathcal{Q}}$ aproxima p
- σ pequeno $\implies p_{\mathcal{Q}} pprox \delta_0$ e $p_\pi pprox p|_{\mathcal{D}} \implies p_\pi \otimes p_{\mathcal{Q}}$ aproxima p
- assim, $I(X_{\pi}; X_{\mathcal{Q}}) = D_{\mathsf{KL}}(p||p_{\pi} \otimes p_{\mathcal{Q}}) \to 0$ para σ grande ou para σ pequeno



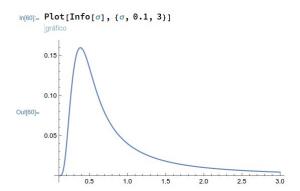


Figura: $I(X_{\pi}; X_{\mathcal{Q}})$ em função de σ .

• máximo: $\sigma \approx 0.38$

Outras perspectivas

• estudar o comportamento da informação de Fisher por enrolamento em quantização ($G_{\pi} \leq G$ e $G_{\mathcal{Q}} \leq G$)



Outras perspectivas

- estudar o comportamento da informação de Fisher por enrolamento em quantização ($G_{\pi} \leq G$ e $G_{\mathcal{O}} \leq G$)
- estudar medidas de distância e divergência nas distribuições enrolada e quantizada (tendo em vista o fator de achatamento)



Outras perspectivas

- estudar o comportamento da informação de Fisher por enrolamento em quantização ($G_{\pi} \leq G$ e $G_{\mathcal{O}} \leq G$)
- estudar medidas de distância e divergência nas distribuições enrolada e quantizada (tendo em vista o fator de achatamento)
- relação com transformada de Fourier: $\widehat{p_{\pi}} = \widehat{p}|_{\Lambda^*}$



- [1] T. M. Cover e Joy A. Thomas. *Elements of information theory*. 2nd ed. OCLC: ocm59879802. Hoboken, N.J: Wiley-Interscience, 2006. ISBN: 9780471241959.
- [2] Cong Ling e Jean-Claude Belfiore. "Achieving AWGN Channel Capacity With Lattice Gaussian Coding". Em: *IEEE Transactions on Information Theory* 60.10 (out. de 2014), pp. 5918–5929. ISSN: 1557-9654. DOI: 10.1109/TIT.2014.2332343.
- [3] Cong Ling, Laura Luzzi e Jean-Claude Belfiore. "Lattice codes achieving strong secrecy over the mod-Λ Gaussian Channel". Em: 2012 IEEE International Symposium on Information Theory Proceedings. ISSN: 2157-8117. Jul. de 2012, pp. 2306–2310. DOI: 10.1109/ISIT.2012.6283924.
- [4] K. V Mardia e Peter E Jupp. *Directional statistics*. English. OCLC: 1039171708. Chichester; New York: J. Wiley, 2010. ISBN: 9780470317815 9780470316979.



Referências