Integration

Cheng Soon Ong

Data61, CSIRO

chengsoon.ong@anu.edu.au

@ChengSoonOng

Marc Peter Deisenroth

University College London

m.deisenroth@ucl.ac.uk

@mpd37

December 2020

Integration problems

- Moment computation

$$M_k(x) = \int x^k p(x) dx$$

Integration problems

- Moment computation

$$M_k(x) = \int x^k p(x) dx$$

- Evidence (marginal likelihood)

$$p(\boldsymbol{X}) = \int p(\boldsymbol{X}|\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

Integration problems

- Moment computation

$$M_k(x) = \int x^k p(x) dx$$

- Evidence (marginal likelihood)

$$p(\boldsymbol{X}) = \int p(\boldsymbol{X}|\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

- Relative entropy (KL divergence)

$$\mathrm{KL}(p||q) = \int \log \frac{p(\boldsymbol{x})}{q(\boldsymbol{x})} p(\boldsymbol{x}) d\boldsymbol{x}$$

Integration problems

- Moment computation

$$M_k(x) = \int x^k p(x) dx$$

- Evidence (marginal likelihood)

$$p(\boldsymbol{X}) = \int p(\boldsymbol{X}|\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

- Relative entropy (KL divergence)

$$\mathsf{KL}(p||q) = \int \log \frac{p(\boldsymbol{x})}{q(\boldsymbol{x})} p(\boldsymbol{x}) d\boldsymbol{x}$$

- Prediction in time-series models

$$p(\boldsymbol{x}_{t+1}) = \int p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t) p(\boldsymbol{x}_t) d\boldsymbol{x}_t$$

Integration problems

- Moment computation

$$M_k(x) = \int x^k p(x) dx$$

- Evidence (marginal likelihood)

$$p(\boldsymbol{X}) = \int p(\boldsymbol{X}|\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

- Relative entropy (KL divergence)

$$\mathsf{KL}(p||q) = \int \log \frac{p(\boldsymbol{x})}{q(\boldsymbol{x})} p(\boldsymbol{x}) d\boldsymbol{x}$$

- Prediction in time-series models

$$p(\boldsymbol{x}_{t+1}) = \int p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t) p(\boldsymbol{x}_t) d\boldsymbol{x}_t$$

- Experimental design

$$\mathbb{E}_{\boldsymbol{\theta}}[U(\boldsymbol{x})] = \int U(\boldsymbol{x}; \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

Integration problems

- Moment computation

$$M_k(x) = \int x^k p(x) dx$$

- Evidence (marginal likelihood)

$$p(\boldsymbol{X}) = \int p(\boldsymbol{X}|\boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

- Relative entropy (KL divergence)

$$\mathsf{KL}(p||q) = \int \log \frac{p(\boldsymbol{x})}{q(\boldsymbol{x})} p(\boldsymbol{x}) d\boldsymbol{x}$$

- Prediction in time-series models

$$p(\boldsymbol{x}_{t+1}) = \int p(\boldsymbol{x}_{t+1}|\boldsymbol{x}_t) p(\boldsymbol{x}_t) d\boldsymbol{x}_t$$

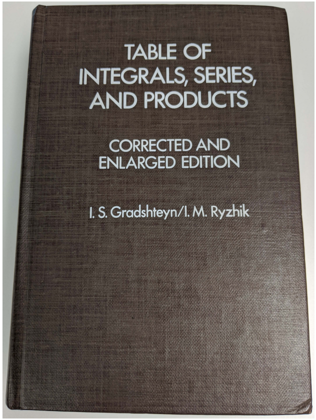
- Experimental design

$$\mathbb{E}_{\boldsymbol{\theta}}[U(\boldsymbol{x})] = \int U(\boldsymbol{x}; \boldsymbol{\theta}) p(\boldsymbol{\theta}) d\boldsymbol{\theta}$$

- Planning

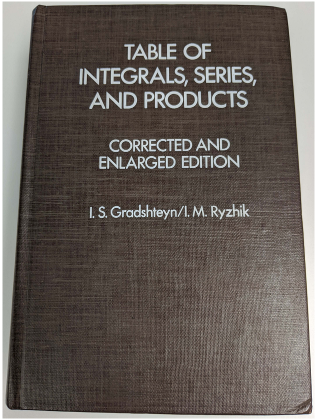
$$J^{\pi}(\boldsymbol{x}(0)) = \int_0^T r(\boldsymbol{x}(t), \boldsymbol{u}(t)) | \boldsymbol{x}(0) dt$$

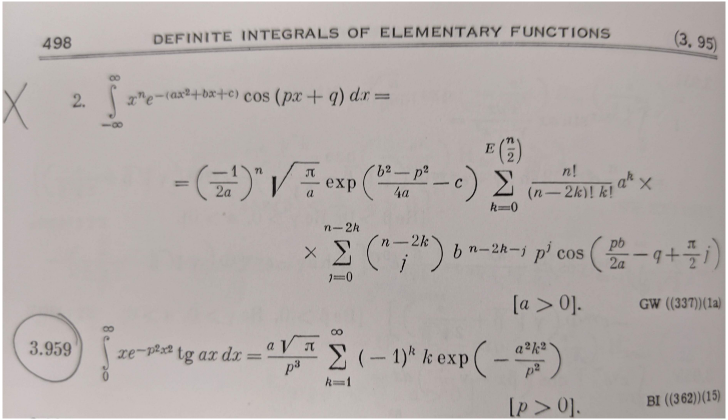
Exact integration



- Compute integrals analytically, if possible (Gradshteyn & Ryzhik, 2007)

Exact integration




498 DEFINITE INTEGRALS OF ELEMENTARY FUNCTIONS (3.95)

2. $$\int_{-\infty}^{\infty} x^n e^{-(ax^2+bx+c)} \cos(px+q)\,dx = \left(\frac{-1}{2a}\right)^n \sqrt{\frac{\pi}{a}} \exp\left(\frac{b^2-p^2}{4a}-c\right) \sum_{k=0}^{E\left(\frac{n}{2}\right)} \frac{n!}{(n-2k)!\,k!} a^k \times \sum_{j=0}^{n-2k} \binom{n-2k}{j} b^{\,n-2k-j} p^j \cos\left(\frac{pb}{2a}-q+\frac{\pi}{2}j\right)$$

$[a>0]$. GW ((337))(1a)

3.959 $$\int_0^\infty x e^{-p^2x^2} \operatorname{tg} ax\,dx = \frac{a\sqrt{\pi}}{p^3} \sum_{k=1}^{\infty} (-1)^k k \exp\left(-\frac{a^2k^2}{p^2}\right)$$

$[p>0]$. BI ((362))(15)

- Compute integrals analytically, if possible (Gradshteyn & Ryzhik, 2007)

Approximate integration

Topic	Useful reference	Video
Numerical integration	Stoer & Bulirsch (2002)	Chapter 1
Bayesian quadrature	Rasmussen & Ghahramani (2003), Gunter et al. (2014)	Chapter 1
Monte-Carlo integration	MacKay (2003), Murray (2015)	Chapter 2
Normalizing flows	Weng (2018), Papamakarios et al. (2019), Kobyzev et al. (2020)	Chapter 3
Inference in time series	Julier & Uhlmann (2004), Särkkä (2013)	Chapter 4

References

Gradshteyn, I. S. and Ryzhik, I. M. (2007). *Table of Integrals, Series, and Products*. Academic Press, 7th edition.

Gunter, T., Osborne, M. A., Garnett, R., Hennig, P., and Roberts, S. J. (2014). Sampling for Inference in Probabilistic Models with Fast Bayesian Quadrature. In *Advances in Neural Information Processing Systems*.

Julier, S. J. and Uhlmann, J. K. (2004). Unscented Filtering and Nonlinear Estimation. *Proceedings of the IEEE*, 92(3):401–422.

Kobyzev, I., Prince, S., and Brubaker, M. (2020). Normalizing Flows: An Introduction and Review of Current Methods. *IEEE Transactions on Pattern Analysis and Machine Intelligence*.

MacKay, D. J. C. (2003). *Information Theory, Inference, and Learning Algorithms*. Cambridge University Press.

Murray, I. (2015). Monte Carlo Inference Methods. *NeurIPS Tutorial*.

References (cont.)

Papamakarios, G., Nalisnick, E., Rezende, D. J., Mohamed, S., and Lakshminarayanan, B. (2019). Normalizing Flows for Probabilistic Modeling and Inference. *arXiv:1912.02762*.

Rasmussen, C. E. and Ghahramani, Z. (2003). Bayesian Monte Carlo. In *Advances in Neural Information Processing Systems*.

Särkkä, S., Solin, A., and Hartikainen, J. (2013). Spatiotemporal Learning via Infinite-Dimensional Bayesian Filtering and Smoothing: A Look at Gaussian Process Regression Through Kalman Filtering,. *IEEE Signal Processing Magazine*, 30(4):51–61.

Stoer, J. and Bulirsch, R. (2002). *Introduction to Numerical Analysis*. Texts in Applied Mathematics. Springer-Verlag, 3rd edition.

Weng, L. (2018). Flow-based Deep Generative Models. *lilianweng.github.io/lil-log*.