MÁSTER EN MATEMÁTICAS, CURSO 2024-25
UNIVERSIDAD AUTÓNOMA DE MADRID

ONDÍCULAS, REDES NEURONALES Y TRATAMIENTO DE SEÑALES
Wavelets, Neural Networks and Signal Processing

Davide Barbieri

Plan del Curso

Cap I - Análisis de Fourier

1. Introducción general al curso. Análisis de Fourier I: la transformada de Fourier. [RS80, Fol92]
2. Análisis de Fourier II: series de Fourier y bases de cosenos. Señales finitas, transformada de Fourier discreta y FFT. [Rud76, Fol92, HW96, Fra99, Kat04]
3. Análisis de Fourier III: teorema de Shannon, espacio de Paley-Wiener y RKHS. [Hig85, Dau92, Fol92]
4. Análisis de Fourier IV: el principio de incertidumbre. [Fol92, DS98, Kat04]
5. *Laboratorio 1*: series temporales, frecuencia de Nyquist, aliasing, convoluciones, módulo y fase.

Cap II – Marcos y análisis en tiempo-frecuencia

6. Minicurso de teoría de marcos. Marcos de Parseval y bases ortonormales. Sucesiones de Bessel, marcos duales y la fórmula de reconstrucción. Maros finitos. [Dau92, HW96, Cas00, Chr16]
7. Análisis en tiempo-frecuencia I: la transformada de Fourier con ventana. Localización en tiempo-frecuencia. [Dau92, Gro01]
8. Análisis en tiempo-frecuencia II: marcos de Gabor. Existencia, el teorema de densidad y el teorema de Balian-Low. [Dau92, HW96, Gro01, Jan05]
9. Análisis en tiempo-frecuencia II: la transformada de Gabor discreta de señales finitas. [Jan94, FS98, Chr16]
10. *Laboratorio 2*: análisis en tiempo-frecuencia de señales audio digitales con marcos de Gabor.

Cap III - Ondículas

11. Introducción a las ondículas I: la transformada de ondículas continua, admisibilidad de Calderón, cancelaciones y detalles. [Dau92, Hol95, Mal09]
12. Introducción a las ondículas II: las ondículas discretas no son sistemas de traslaciones. Ondículas de Haar y de Shannon. [HW96, GLWW03, Mal09]
13. Análisis en Multirresolución I: aproximación, detalle y la función de escala. [Dau92, HW96, Mal09]
14. Análisis en Multirresolución II: el teorema de completitud y el filtro paso bajo de una MRA. [HW96]
15. Análisis en Multirresolución III: el filtro de paso alto y las MRA ortonormales. [Dau92, HW96, Mal09]
16. Análisis en Multirresolución IV: ondículas de soporte compacto y el algoritmo de descomposición. [Dau92, HW96, Mal09, Pin02]
17. *Laboratorio 3*: las ondículas de Morlet y de sombrero Mexicano; análisis MRA de imágenes digitales. [Fra99, VFl08]

Cap IV – Aprendizaje automático y redes neuronales

18. Introducción al aprendizaje automático: aprendizaje supervisado, no supervisado y por refuerzo.
y arquitecturas básicas: la neurona, el MLP, las redes convolucionales.
19. Introducción a las redes neuronales: arquitecturas básicas y descenso del gradiente. [Hay09, Nes03]
20. Más sobre el descenso del gradiente: backpropagation y online learning. [Hay09, Zin03, Haz22]
21. Redes neuronales y teoría de aproximación I: el teorema de aproximación universal de Cybenko. [Pin99]
22. Redes neuronales y teoría de aproximación II: aproximación profunda. [AB09, Pog17, Yar17]
23. Cuestiones abiertas sobre redes neuronales: el problema de bias-variance y el exceso de parametrización, regularización implícita en la SGD. [CS01, CS18, BHM19]
24. Minicurso de Reinforcement Learning, I. Procesos de decisión de Markov y Dynamic Programming. [Put14]
25. Minicurso de Reinforcement Learning, II. El Policy Gradient Theorem y la Proximal Policy Optimization. [SB18, OAI17]

Bibliografía

Cap I

[Bre11] H. Brezis, Functional analysis, Sobolev spaces and partial differential equations. Springer, 2011.
[DS98] D. Donoho, P. Stark, Uncertainty principles and signal recovery. SIAM J. Appl. Math. 49, pp. 906-931, 1998.
[Con90] J. Conway, A course in functional analysis. Springer, 2nd ed. 1990.
[Fol92] G. Folland, Fourier analysis and its applications. Brooks/Cole, 1992.
[Hig85] J. Higgins, Five short stories about the cardinal series. Bull. Amer. Math. Soc. (N.S.) 12, pp. 45-89, 1985.
[Kat04] Y. Katznelson, An introduction to harmonic analysis. Cambridge University Press, 3rd ed. 2004.
[RS80] M. Reed, B. Simon, Methods of modern mathematical physics Vol. 1 - Functional analysis. Academic Press 1980.
[Rud76] W. Rudin, Principles of mathematical analysis. McGraw-Hill, 3rd ed. 1976.

Cap II y III

[Cas00] P. Casazza, The art of frame theory. Taiwanese J. Math. 4, pp.129-201, 2000.
[Chr16] O. Christensen. An introduction to frames and Riesz bases. Birkhäuser, 2nd ed. 2016.
[Dau92] I. Daubechies. Ten lectures on wavelets. SIAM, 1992.
[FS98] H. Feichtinger, T. Strohmer, Eds. Gabor analysis and algorithms. Springer, 1998.
[Fra99] M. Frazier. An introduction to wavelets through linear algebra. Springer, 1999.
[GLWW03] P. Gressman, D. Labate, G. Weiss, E. Wilson, Affine, quasi-affine and co-affine wavelets. In "Beyond wavelets", G. Welland (Ed.), Elsevier 2003.
[Gro01] K. Gröchenig. Foundations of time-frequency analysis. Springer, 2001.
[HW96] E. Hernández, G. Weiss. A first course on wavelets. CRC Press, 1996.
[Hol95] M. Holschneider. Wavelets. An analysis tool. Clarendon Press 1995.
[Jan94] A. Janssen, Duality and biorthogonality for discrete-time Weyl-Heisenberg frames. Unclassified report, Philips Electronics, 002/94.
[Jan05] A. Janssen, Classroom proof of the density theorem for Gabor systems. ESI Preprint, 2005.
[Mal09] S. Mallat. A wavelet tour of signal processing. Academic Press, 3rd ed. 2009.
[Pin02] M. A. Pinsky. Introduction to Fourier analysis and wavelets. AMS 2002.
[SN97] G. Strang, T. Nguyen. Wavelets and filter banks. Wellesley-Cambridge Press, 1997.
[VFl08] P. Van Fleet. Discrete wavelet transformations. Wiley, 2008.

Cap IV

[AB09] M. Anthony and P.L. Barlett. Neuronal network learning: Theoretical foundations. Cambridge University Press, 2009.
[BHM19] M. Belkin, D. Hsu, S. Mandal. Reconciling modern machine-learning practice and the classical bias–variance trade-off. PNAS 116:15849-15854 (2019).
[CS18] P. Chaudari, S. Soatto. Stochastic Gradient Descent Performs Variational Inference, Converges to Limit Cycles for Deep Networks. Information Theory and Applications Workshop ITA (2018).
[CS01] F. Cucker, S. Smale. On the Mathematical Foundations of Learning. Bulletin of the AMS 39:1-49 (2001).
[Hay09] S. Haykin. Neural Networks and Learning Machines. Pearson, 3rd ed. 2009.
[Haz22] E. Hazan. Introduction to online convex optimization. MIT Press, 2nd ed. 2022.
[LBRN06] H. Lee, A. Battle, R. Raina & A. Y. Ng. Efficient sparse coding algorithms. NIPS (2006).
[vL07] U. von Luxburg. A tutorial on spectral clustering. Stat Comput 17: 395–416 (2007).
[Nes03] Y. Nesterov. Introductory lectures on convex optimization. Springer, 2003.
[OAI17] J. Schulman, F. Wolski, P. Dhariwal, A. Radford, O. Klimov. Proximal Policy Optimization Algorithms. OpenAI preprint, arXiv.
[Pin99] A. Pinkus. Approximation theory of the MLP model in neural networks, Acta Numerica 8:143-195 (1999).
[Pog17] T. Poggio et al. Why and When Can Deep-but Not Shallow-networks Avoid the Curse of Dimensionality: A Review. International Journal of Automation and Computing 14:503-519 (2017)
[Put14] M. L. Puterman. Markov decision processes: discrete stochastic dynamic programming, John Wiley & Sons, 2014.
[SB18] R. S. Sutton and A. G. Barto. Reinforcement learning: An introduction. MIT press, 2018.
[Yar17] D. Yarotsky. Error bounds for approximations with deep ReLU networks. Neural Networks, 94:103-114 (2017).
[Zin03] M. Zinkevich. Online convex programming and generalized infinitesimal gradient ascent. ICML (2003).

Descripción

La propuesta está estructurada en 4 partes:

I) Fundamentos del análisis de Fourier. Se introducen las nociones fundamentales sobre suavidad en relación a las oscilaciones, muestreo, principio de incertidumbre, señales finitas y digitales. Sobre estas últimas se hace un laboratorio.

II, III) Análisis de las señales en tiempo-frecuencia y en tiempo-escala. Se introducen las dos familias más conocidas de características locales para el estudio de las señales: la frecuencia local, y la escala local. Se estudian por separado sus propiedades en la teoría continua y en la teoría discreta, también con laboratorios.

IV) Aprendizaje automático. Se introducen las redes neuronales como funciones para problemas de aproximación. Se discuten aspectos numéricos de optimización y aspectos estadísticos de rigidez al respecto del overfitting. Se muestran algunas arquitecturas relacionadas con las partes anteriores, en particular las redes convolucionales. En la parte final se introduce una técnica de aprendizaje con la que se entrenan las redes modernas.

El objetivo del curso es proporcionar los conceptos matemáticos básicos para poder tener un enfoque cuantitativo hacia los aspectos más en uso del análisis de las señales. Las partes I-III) de esta asignatura están muy relacionadas con análisis real, análisis de Fourier y análisis funcional, pero están diseñadas para poderse seguir de forma más o menos independiente de esas asignaturas. La parte IV) está muy relacionada no solo con lo que se hace anteriormente en esta asignatura, sino también con los cursos de estadística, pero aquí también se intenta dar todas las nociones necesarias en la asignatura.

Los laboratorios tienen un carácter principalmente ilustrativo: no se pretende que los estudiantes programen. Pero por experiencia es muy útil ver en práctica la implementación de los conceptos. No se plantea un laboratorio para la parte IV porque ya sería algo que requiere más que una clase (y probablemente más que dos). Se propondrá usar herramientas de aprendizaje automático (Tensorflow en Google Colab) para las presentaciones finales.

Evaluación:

- 2 entregas de teoría, sobre el contenido de las partes I-III): 40% de la nota
- Presentación final de 30 minutos: 60% de la nota.