

Separating Out the Eigenvalue Densities: Computing the Jacobians

Random Matrix Theory with Its Applications

Leyi Zhao

School of Cyberspace

2024 年 4 月 2 日



- ① Background
- ② A Two-Dimensional Example
- ③ 监督学习

① Background

② A Two-Dimensional Example

③ 监督学习

Spectral Theorem

- The spectral theorem:

$$M = U\Lambda U^* (\text{unitary, symplectic})$$

$$M = U\Lambda U^T (\text{orthogonal})$$

M can be any kinds of matrices.

- Three ensembles:¹
 - Gaussian Orthogonal Ensemble**: the set of $N \times N$ random **real symmetric** matrices
 - Gaussian Unitary Ensemble**: the set of $N \times N$ random **complex Hermitian** matrices
 - Gaussian Symplectic Ensemble**: the set of $N \times N$ random **quaternion (四元数) self-dual Hermitian** matrices
- 这可以被视为一种坐标系变化: $M \mapsto (\Lambda, U)$
- 此时可以考虑计算雅可比行列式 (Jacobian): $\left| \det \frac{\partial M}{\partial (\Lambda, U)} \right|$

¹Yanqing Yin and Zhidong Bai. "Convergence rates of the spectral distributions of large random quaternion self-dual Hermitian matrices". In: *Journal of Statistical Physics* 157 (2014), pp. 1207–1224.

Gaussian Orthogonal Ensemble(GOE)

- **Gaussian Orthogonal Ensemble:** the set of $N \times N$ random **real symmetric** matrices
- Symmetric matrix:
 - $S = S^T$
 - $S = X\Lambda X^{-1}, S^T = (X^{-1})^T \Lambda X^T$ when $S = S^T, X^T X = I$
 - A symmetric matrix has only *real eigenvalues*.
 - Symmetric diagonalization:
 $S = Q\Lambda Q^{-1} = Q\Lambda Q^T$ with $Q^{-1} = Q^T$
 - Example: $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$
 - Free variables number is $n + \frac{n(n-1)}{2} \cdot 1$

Gaussian Unitary Ensemble(GUE)

- **Gaussian Unitary Ensemble**: the set of $N \times N$ random **complex Hermitian** matrices
- Hermitian matrices:
 - $S = S^*$
 - $S = U\Lambda U^{-1}, S^* = (U^{-1})^* \Lambda U^*$ when $S = S^*, U^* U = I$
 - Hermitian diagonalization:
 $S = U\Lambda U^{-1} = U\Lambda U^*$ with $U^{-1} = U^*$
 - Example:
$$\begin{bmatrix} a & b + ci \\ b - ci & d \end{bmatrix}$$
 - Free variables number is $n + \frac{n(n-1)}{2} \cdot 2$

Gaussian Symplectic Ensemble(GSE)

- **Gaussian Symplectic Ensemble**: the set of $N \times N$ random **quaternion self-dual Hermitian** matrices⁴⁵
- Quaternions were invented in 1843 by the Irish mathematician William Rowen Hamilton as an extension of complex numbers into three dimensions⁶⁷, and it's well known that the quaternion field \mathbb{Q} can be represented as a two-dimensional complex vector space⁸.

⁴许方官. 四元数物理学. 北京大学出版社, 2012.

⁵Yanqing Yin and Zhidong Bai. "Convergence rates of the spectral distributions of large random quaternion self-dual Hermitian matrices". In: *Journal of Statistical Physics* 157 (2014), pp. 1207–1224.

⁶Cipher A Deavours. "The quaternion calculus". In: *The American Mathematical Monthly* 80.9 (1973), pp. 995–1008.

⁷William Rowan Hamilton. *Elements of quaternions*. London: Longmans, Green, & Company, 1866.

⁸Claude Chevalley. *Theory of Lie Groups (PMS-8)*. 1946.

四元数

- 一个三位空间的实矢量 $\mathbf{a} = a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k}$ 与一个实数 a_0 组合成一个数 $A = a_0 + \mathbf{a}$ 被称为**四元数**。其中实数 a_0 被称为四元数 A 的**标部**，实矢量 \mathbf{a} 称为四元数 A 的**矢部**。
- 在四元数 A 的矢部中，三个矢量 $\mathbf{i}, \mathbf{j}, \mathbf{k}$ 是三维空间中三个互相垂直且方向固定的单位矢量。
- 一个四元数实际上是由四个基 $1, \mathbf{i}, \mathbf{j}, \mathbf{k}$ 的线性组合构成的。

四元数代数运算规则

- **相等**: 两个四元数 A 和 B , 若他们相应的四个分量分别相等, 即

$$a_n = b_n, \quad n = 0, 1, 2, 3$$

则称他们**相等**, 记作

$$A = B$$

一个四元数的等式相当于四个实数等式。

- **加法**: 两个四元数 A 和 B 之和 C 仍为四元数, 记作

$$C = A + B$$

其定义为

$$c_n = a_n + b_n, \quad n = 0, 1, 2, 3$$

由实数加法的交换律、结合律知, 四元数加法也存在着交换律和结合律。

四元数代数运算规则

- 乘法: 两个四元数 A 和 B 乘积的定义是

$$\begin{aligned}AB &= (a_0 + \mathbf{a})(b_0 + \mathbf{b}) \\&= a_0 b_0 - \mathbf{a} \cdot \mathbf{b} + a_0 \mathbf{b} + b_0 \mathbf{a} + \mathbf{a} \times \mathbf{b} \\BA &= (b_0 + \mathbf{b})(a_0 + \mathbf{a}) \\&= a_0 b_0 - \mathbf{a} \cdot \mathbf{b} + a_0 \mathbf{b} + b_0 \mathbf{a} + \mathbf{b} \times \mathbf{a}\end{aligned}$$

两个四元数的乘积仍是四元数, 但因为 $\mathbf{a} \times \mathbf{b} \neq \mathbf{b} \times \mathbf{a}$, 故
 $AB \neq BA$

四元数的乘法规则是由三个互相垂直的空间单位矢量 $\mathbf{i}, \mathbf{j}, \mathbf{k}$ 之间的乘法 (叉乘) 规则

$$\begin{cases} \mathbf{i}\mathbf{i} = \mathbf{j}\mathbf{j} = \mathbf{k}\mathbf{k} = -1 \\ \mathbf{i}\mathbf{j} = -\mathbf{j}\mathbf{i} = \mathbf{k} \\ \mathbf{j}\mathbf{k} = -\mathbf{k}\mathbf{j} = \mathbf{i} \\ \mathbf{k}\mathbf{i} = -\mathbf{i}\mathbf{k} = \mathbf{j} \end{cases}$$

所导致的

四元数代数运算规则

- **共轭 (四元共轭):** 与四元数 $A = a_0 + \mathbf{a}$ 对应的另一个四元数

$$\bar{A} = a_0 - \mathbf{a}$$

称为 A 的**四元共轭**, 反之, A 也称为 \bar{A} 的**四元共轭**
此外, 由

$$\overline{AB} = a_0 b_0 - \mathbf{a} \cdot \mathbf{b} - a_0 \mathbf{b} - b_0 \mathbf{a} - \mathbf{a} \times \mathbf{b}$$

$$\begin{aligned}\overline{B\bar{A}} &= (b_0 - \mathbf{b})(a_0 - \mathbf{a}) \\ &= a_0 b_0 - \mathbf{a} \cdot \mathbf{b} - a_0 \mathbf{b} - b_0 \mathbf{a} + \mathbf{b} \times \mathbf{a}\end{aligned}$$

可知

$$\overline{AB} = \overline{B\bar{A}}$$

- **模方:** 称实数

$$\|A\| = \|\bar{A}\| = A\bar{A} = \bar{A}A = a_0^2 + \mathbf{a} \cdot \mathbf{a}$$

为四元数 A 或 \bar{A} 的**模方**, 且两个四元数之积的模方等于两个四元数模方之积
 $\|AB\| = \|A\| \|B\|$

四元数的矩阵表示

- **一级四元数与 2×2 矩阵:** 把四元数的四个基 $1, i, j, k$ 分别用四个 2×2 矩阵表示成

$$1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}, \quad i = \begin{bmatrix} i & 0 \\ 0 & -i \end{bmatrix}, \quad j = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}, \quad k = \begin{bmatrix} 0 & i \\ i & 0 \end{bmatrix}$$

这样，一个四元数可以用一个 2×2 的复矩阵表示成

$$A = a_0 + \mathbf{a} = \begin{bmatrix} a_0 + ia_1 & a_2 + ia_3 \\ -a_2 + ia_3 & a_0 - ia_1 \end{bmatrix} = \begin{bmatrix} \alpha & \beta \\ -\bar{\beta} & \bar{\alpha} \end{bmatrix}$$

- **共轭:**

$$\bar{A} = a_0 - a_1 i - a_2 j - a_3 k = \begin{bmatrix} a_0 - ia_1 & -a_2 - ia_3 \\ a_2 - ia_3 & a_0 + ia_1 \end{bmatrix} = \begin{bmatrix} \bar{\alpha} & -\beta \\ \bar{\beta} & \alpha \end{bmatrix}$$

- **模方:** 四元数的模方就是其表示的矩阵的行列式，即

$$\|A\| = \det A = \sqrt{a_0^2 + a_1^2 + a_2^2 + a_3^2} = \sqrt{|\alpha|^2 + |\beta|^2}$$

四元数的矩阵表示

一个四元数可以用 2×2 矩阵表示, 类似地, 二级四元数可以用 4×4 矩阵表示, 其四个基 $1, i_2, j_2, k_2$ 表示成

$$1 = \begin{bmatrix} I_2 & 0 \\ 0 & I_2 \end{bmatrix}, \quad i = \begin{bmatrix} iI_2 & 0 \\ 0 & -iI_2 \end{bmatrix}, \quad j = \begin{bmatrix} 0 & I_2 \\ -I_2 & 0 \end{bmatrix}, \quad k = \begin{bmatrix} 0 & iI_2 \\ iI_2 & 0 \end{bmatrix}$$

其中 I_2 是 2×2 的单位矩阵。这样, 一个二级四元数的矩阵表示为

$$\begin{aligned} Q &= A + Bi_2 + Cj_2 + Dk_2 \\ &= \begin{bmatrix} A + iB & C + iD \\ -C + iD & A - iB \end{bmatrix} \\ &= \begin{bmatrix} a_0 - b_1 + i(a_1 + b_0) & a_2 - b_3 + i(a_3 + b_2) & c_0 - d_1 + i(c_1 + d_0) & c_2 - d_3 + i(c_3 + d_2) \\ -(a_2 + b_3) + i(a_3 - b_2) & a_0 + b_1 - i(a_1 - b_0) & -(c_2 + d_3) + i(c_3 - d_2) & c_0 + d_1 - i(c_1 - d_0) \\ -(c_0 + d_1) - i(c_1 - d_0) & -(c_2 + d_3) - i(c_3 - d_2) & a_0 + b_1 + i(a_1 - b_0) & a_2 + b_3 + i(a_3 - b_2) \\ c_2 - d_3 - i(c_3 + d_2) & -(c_0 - d_1) + i(c_1 + d_0) & -(a_2 - b_3) + i(a_3 + b_2) & a_0 - b_1 - i(a_1 + b_0) \end{bmatrix} \end{aligned}$$

Gaussian Symplectic Ensemble(GSE)

- An $n \times n$ **quaternion self-dual Hermitian matrix** ${}^q\mathbf{A}_n = (x_{jk})_{n \times n}$ is a matrix whose entries $x_j(j, k = 1, \dots, n)$ are **quaternions** and satisfy $x_{jk} = \bar{x}_{kj}$. If each x_{jk} is viewed as a 2×2 matrix, then $\mathbf{A}_n = (x_{jk})$ is in fact a $2n \times 2n$ Hermitian matrix. And, the entries of \mathbf{A}_n can be represented as

$$x_{jk} = \begin{bmatrix} a_{jk} + b_{jk}i & c_{jk} + d_{jk}i \\ -c_{jk} + d_{jk}i & a_{jk} - b_{jk}i \end{bmatrix} = \begin{bmatrix} \alpha_{jk} & \beta_{jk} \\ -\bar{\beta}_{jk} & \bar{\alpha}_{jk} \end{bmatrix}, 1 \leq j < k \leq n,$$

and $x_{jj} = \begin{bmatrix} a_{jj} & 0 \\ 0 & a_{jj} \end{bmatrix}$, where $a_{jk}, b_{jk}, c_{jk}, d_{jk} \in \mathbb{R}$ and $1 \leq j, k \leq n$. It is well known that the multiplicities of all the eigenvalues of \mathcal{A}_n are even⁹.

- Example: $\begin{bmatrix} \begin{bmatrix} a_{11} & 0 \\ 0 & a_{11} \end{bmatrix} & \begin{bmatrix} a_{12} + b_{12}i & c_{12} + d_{12}i \\ -c_{12} + d_{12}i & a_{12} - b_{12}i \end{bmatrix} \\ \begin{bmatrix} a_{12} - b_{12}i & -c_{12} - d_{12}i \\ c_{12} - d_{12}i & a_{12} + b_{12}i \end{bmatrix} & \begin{bmatrix} a_{44} & 0 \\ 0 & a_{44} \end{bmatrix} \end{bmatrix}$
- Free variables number is $n + \frac{n(n-1)}{2} \cdot 4$

⁹Fuzhen Zhang. "Quaternions and matrices of quaternions". In: *Linear algebra and its applications* 251 (1997), pp. 21–57.

① Background

② A Two-Dimensional Example

③ 监督学习

2×2 Real Symmetric Matrix

- Consider a 2×2 real symmetric matrix $M = \begin{bmatrix} a & b \\ b & c \end{bmatrix}$ with distinct eigenvalues $\lambda_1 \neq \lambda_2$

① Background

② A Two-Dimensional Example

③ 监督学习

监督学习


- 监督学习，作为机器学习的核心范式之一，主要依赖于充分标记的数据集来训练算法。
- 常见的技术：支持向量机、决策树、随机森林、朴素贝叶斯、K 近邻算法、神经网络、梯度提升、线性回归、逻辑回归
- 简述四种下游任务：
 - 文本分类
 - 图像识别
 - 机器翻译
 - 异常检测

文本分类

- 文本分类是一项自动将文本文档分类到预定义类别的任务¹⁰
- 支持向量机：SVMs 构建在文档类别之间的最优分隔超平面上，以分类新的未标记示例¹¹
- k-最近邻：kNN 算法根据新文档与训练集中标记示例的相似性/距离来预测类别¹²

¹⁰Femi Joseph and N Ramakrishnan. "Text categorization using improved K nearest neighbor algorithm". In: *Int J Trends Eng Technol* 4 (2015), pp. 65–68, Krystian Horecki and Jacek Mazurkiewicz. "Natural language processing methods used for automatic prediction mechanism of related phenomenon". In: *International Conference on Artificial Intelligence and Soft Computing*. Springer. 2015, pp. 13–24.

¹¹Khushbu Khamar. "Short text classification using kNN based on distance function". In: *International Journal of advanced research in computer and communication engineering* 2.4 (2013), pp. 1916–1919.

¹²Kwok James Tin-Yau. "Automated text categorization using support vector machine". In: *Proc. the International Conference on Neural Information Processing*. 1999, pp. 347–351, Jason DM Rennie and Ryan Rifkin. "Improving multiclass text classification with the support vector machine". In: (2001). 

图像识别

- 经典的特征提取与分类器方法
- 卷积神经网络方法 (CNN)

经典的特征提取与分类器方法

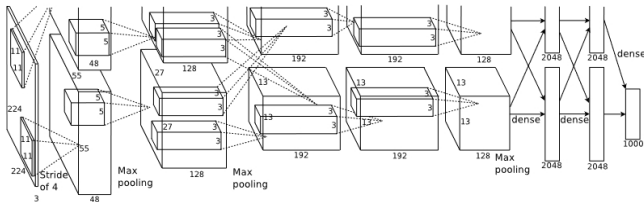
- 在 20 世纪 90 年代之前传统的机器学习时代，主要使用允许特征工程和传统机器学习算法来进行图像识别。这一时期的研究者和工程师们需要大量依赖专业知识和领域经验，去手工设计识别各类目标所需的特征提取算法。这类特征通常针对不同图像识别任务定制，主要有 SIFT¹³, HOG¹⁴等。

¹³David G Lowe. "Object recognition from local scale-invariant features". In: *Proceedings of the seventh IEEE international conference on computer vision*. Vol. 2. Ieee. 1999, pp. 1150–1157.

¹⁴N. Dalal and B. Triggs. "Histograms of oriented gradients for human detection". In: *2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05)*. Vol. 1. 2005, 886–893 vol. 1. DOI: 10.1109/CVPR.2005.177.

卷积神经网络方法 (CNN)

- AlexNet 在深度学习模型中应用了比较深的 8 卷积层的神经网络, 包含了卷积层, 池化层, 全连接层等模块, 如图所示。这种网络结构的设计成为 续深度网络的范式。之后便发展出了更多使用多层感知机 (MLP), 卷积神经网络 (CNN) 等的模型来进行端到端的图像特征学习和分类¹⁵。

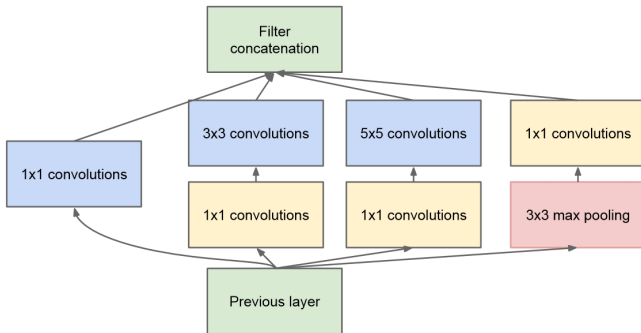


- AlexNet 的成功启示了后续深度学习模型对于层次化特征学习的重要性, 并为图像识别任务提供了一种强大的架构范例。

¹⁵ Alex Krizhevsky, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks". In: *Commun. ACM* 60.6 (May 2017), pp. 84–90. ISSN: 0001-0782. DOI: 10.1145/3065386. URL: <https://doi.org/10.1145/3065386>.

卷积神经网络方法 (CNN)

- GoogLeNet图像识别中采用了独特的 Inception 模块以及一系列创新性的设计，其处理过程充分利用了多尺度特征的丰富性。总体而言，GoogLeNet 是一种注重了模型的深度、宽度和计算效率的模型¹⁶。



¹⁶Christian Szegedy et al. "Going deeper with convolutions". In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2015, pp. 1–9.

机器翻译

- 基于规则的机器翻译 (RBMT)
- 统计机器翻译 (SMT)
- 神经机器翻译 (NMT)
- 度强化学习机器翻译 (DRLMT)

基于规则的机器翻译 (RBMT)

- RBMT 是早期机器翻译方法之一，它的发展可以追溯到 20 世纪 60 年代和 70 年代。如 Fakhrahmad 等人¹⁷于 2012 年针对词义消歧 (WSD) 这一机器翻译过程中最具挑战性的任务，提 各种监督和无监督学习方法来解决这一问题。近年最新的研究包括 Chauhan 等人¹⁸提出的基于规则的模糊计算方法在印地语机器翻译中的自监督情感极性分类和词义消歧方面的应用

¹⁷ Seyed Mostafa Fakhrahmad et al. "A new fuzzy rule-based classification system for word sense disambiguation". In: *Intelligent Data Analysis* 16.4 (2012), pp. 633–648.

¹⁸ Shweta Chauhan et al. "Rule Based Fuzzy Computing Approach on Self-Supervised Sentiment Polarity Classification with Word Sense Disambiguation in Machine Translation for Hindi Language". In: *ACM Transactions on Asian and Low-Resource Language Information Processing* 22.5 (2023), pp. 1–21. 

统计机器翻译 (SMT)

- 基于监督学习的统计机器翻译 (SMT) 是一种传统的机器翻译方法，它使用大量的双语平行语料来训练模型，包括源语言和目标语言之间的句子对。
- González-Rubio 等人¹⁹提出了一个用于计算机辅助翻译的成本敏感主动学习框架，其目标是使翻译过程尽可能轻松。与传统的主动学习场景不同，该论文所提出的主动学习框架的设计不仅是为了最小化用户必须监督的翻译数量，还包括最小化每个翻译的监督难度。

¹⁹Jesús González-Rubio and Francisco Casacuberta. "Cost-sensitive active learning for computer-assisted translation". In: *Pattern Recognition Letters* 37 (2014), pp. 124–134.

异常检测

- 传统监督学习方法
- 深度学习方法
- 应用领域

传统监督学习方法

- 早在 2009 年, Babenko 等人提出了一种基于“多实例学习 (Multiple Instance Learning, MIL)”的方法, 该方法具有潜力应用于异常检测任务²²
- Babenko 等人的研究提出了一种具有实时性能的多实例学习算法, 用于目标跟踪²³。

²²Boris Babenko, Ming-Hsuan Yang, and Serge Belongie. “Visual tracking with online Multiple Instance Learning”. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. June 2009, pp. 983–990. DOI: 10.1109/CVPR.2009.5206737.

²³Boris Babenko, Ming-Hsuan Yang, and Serge Belongie. “Visual tracking with online Multiple Instance Learning”. In: *2009 IEEE Conference on Computer Vision and Pattern Recognition*. June 2009, pp. 983–990. DOI: 10.1109/CVPR.2009.5206737.

深度学习方法

- 一项名为"Deep Learning Anomaly Detection Method in Textual Data" 的研究, 由 Amir Jafari 于 2022 年提出, 探讨了如何结合深度学习和传统机器学习算法来检测和识别文本中的异常²⁴。该研究利用深度学习模型和 Transformer 架构, 将文本数据转化为数值表示, 提供了关于文本数据的关键上下文信息。

²⁴ Amir Jafari. A Deep Learning Anomaly Detection Method in Textual Data. 2022. [arXiv: 2211.13905 \[cs.CL\]](#). ↻ 🔍 🔗

应用领域

- 在 2017 年举行的 KDD Workshop on Anomaly Detection in Finance 中，研究人员和从业者 聚在一起，讨论了这些新方法和解决方案²⁵。
- 一篇题为"Deep Learning based pipeline for anomaly detection and quality enhancement in industrial binder jetting processes" 的研究 由 Alexander Zeiser 等人于 2022 年提，探讨了在工 制 使用深度学习进行异 检测和质量提升的方 ²⁶。
- 一篇题 "Cybersecurity Vital Signs: The Role of Anomaly Detection on Insider Threat Triage" 的研究 由 Karla Clarke 和 Yair Levy2019 年，探讨 异常检测在 部威胁检 中 要作用²⁷。
- 一项名为"Medical Healthcare System Based on Wireless Body Area Networks: The Importance of Anomaly Detection", 由 Hayder Hassaballah、Rashid Fayadh 和 Bushra AlHayali 于 2020 年提出 在 在医疗 WBSN统中 用 检测的现代 架²⁸

²⁵ Archana Anandakrishnan et al. "Anomaly Detection in Finance: Editors' Introduction". In: *Proceedings of the KDD 2017 Workshop on Anomaly Detection in Finance, ADF@KDD 2017, Halifax, Nova Scotia, Canada, August 14, 2017*. Ed. by Archana Anandakrishnan et al. Vol. 71. Proceedings of Machine Learning Research. PMLR, 2017, pp. 1–7. URL: <http://proceedings.mlr.press/v71/anandakrishnan18a.html>.

²⁶Alexander Zeiser, Bas van Stein, and Thomas Bäck. *Deep Learning based pipeline for anomaly detection and quality enhancement in industrial binder jetting processes*. 2022. [arXiv: 2209.10178 \[cs.LG\]](https://arxiv.org/abs/2209.10178).

²⁷Karla Clarke and Yair Levy. "Cybersecurity Vital Signs : The Role of Anomaly Detection on Insider Threat Triage". In: 2019. URL: <https://api.semanticscholar.org/CorpusID:197634071>.

²⁸ Hayder Hassaballah, Rashid Fayadh, and Bushra AlHayali. "Medical Healthcare System Based on Wireless Body Area Networks: The Importance of Anomaly Detection". In: *International Journal of Engineering and Advanced Technology* 9 (Feb. 2020), pp. 3039–3046. DOI: 10.35940/ijeat.C5945.029320.  

Thanks!