Eigenwert

Autor und Herausgeber

Alexander Jung*, Department of Computer Science, Aalto University, Espoo, Finland (ORCID 0000-0001-7538-0990)
Editors: Konstantina Olioumtsevits (ORCID 0000-0003-2697-8659), Ekkehard Schnoor (ORCID 0000-0001-8252-1499), Aalto University, Espoo, Finland
*Corresponding author: alex.jung@aalto.fi

Zusammenfassung

Ein Eigenwert einer quadratischen Matrix $\mathbf{A}$ ist eine Zahl λ, für die ein nicht-null vector $\mathbf{x}$ die Gleichung $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ erfüllt. Ein solcher vector ist ein eigenvector von $\mathbf{A}$. Die Multiplikation eines eigenvectors mit der Matrix skaliert ihn, ohne ihn von der von ihm aufgespannten Geraden zu verschieben. Die Eigenwerte einer diagonalizablen Matrix bilden den diagonalen Faktor ihrer Eigenwertzerlegung. Ein iteratives Verfahren, das eine affine Aktualisierung anwendet, wie Gradientenabstieg für lineare Regression, konvergiert gegen den eindeutigen fixed point der Aktualisierung, für jede Initialisierung, genau dann, wenn jeder Eigenwert der Aktualisierungs-Matrix einen Betrag kleiner als eins hat. Der zweitkleinste Eigenwert der Laplacian matrix eines undirected graphen misst, wie gut der graph connected ist, und ist die Grundlage für spectral clustering.

Definition

Ob Gradientenabstieg für lineare Regression konvergiert, wird durch die Eigenwerte einer einzelnen quadratischen Matrix entschieden. Eine quadratische Matrix $\mathbf{A}$ repräsentiert eine linear map von $\mathbb{R}^d$ nach $\mathbb{R}^d$; das Bild eines vectors $\mathbf{x}$ unter dieser Abbildung ist das Produkt $\mathbf{A}\mathbf{x}$. Die Multiplikation eines vectors mit einer quadratischen Matrix verschiebt ihn normalerweise von der von ihm aufgespannten Geraden. Einige nicht-null vectors sind Ausnahmen: jeder bleibt auf der von ihm aufgespannten Geraden und wird nur skaliert. Der Faktor, um den ein solcher vector skaliert wird, ist ein Eigenwert.

Formal ist eine Zahl $\lambda \in \mathbb{R}$ ein Eigenwert einer quadratischen Matrix $\mathbf{A} \in \mathbb{R}^{d\times d}$, eine mit d Zeilen und d Spalten, wenn $\mathbf{A}\mathbf{x} = \lambda\mathbf{x}$ für einen nicht-null vector $\mathbf{x} \in \mathbb{R}^d$ gilt; ein solcher vector ist ein eigenvector von $\mathbf{A}$ (Strang, 2016, Sect. 6.1). Für einen eigenvector bleibt das Bild $\mathbf{A}\mathbf{x}$ auf der von $\mathbf{x}$ aufgespannten Geraden und zeigt in die gleiche Richtung, wenn $\lambda > 0$, und in die entgegengesetzte Richtung, wenn $\lambda < 0$ (siehe Abb. 1). Das Anordnen von d linearly independenten eigenvectors einer diagonalizablen Matrix als Spalten von $\mathbf{V}$ und der entsprechenden Eigenwerte in einer diagonalen Matrix $\mathbf{\Lambda}$ ergibt die Eigenwertzerlegung $\mathbf{A} = \mathbf{V}\mathbf{\Lambda}\mathbf{V}^{-1}$ (Strang, 2016, Sect. 6.2).

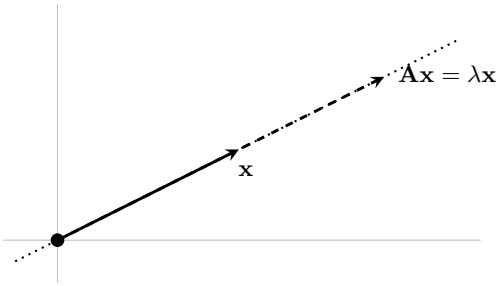


Abbildung 1: Ein eigenvector $\mathbf{x}$ von $\mathbf{A}$ (durchgezogener Pfeil) und sein Bild $\mathbf{Ax} = \lambda\mathbf{x}$ (gestrichelter Pfeil), hier mit Eigenwert $\lambda = 1.8$. Das Bild bleibt auf der gepunkteten Linie, die $\mathbf{x}$ aufspannt, skaliert um λ. Ein negatives λ würde das Bild in die andere Richtung entlang derselben Linie zeigen lassen.

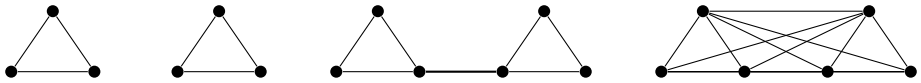
Eigenwerte können verwendet werden, um iterative machine learning (ML)-Verfahren mit einer Aktualisierung der Form

$$\mathbf{w}^{(t+1)} = \mathbf{M}\mathbf{w}^{(t)} + \mathbf{b}, \text{ für } t = 0, 1, \ldots.$$

Ein Beispiel ist Gradientenabstieg für lineare Regression, dessen Aktualisierungsoperator $\mathcal{F}^{(\eta)} : \mathbf{w} \mapsto \mathbf{M}\mathbf{w} + \mathbf{b}$ affin ist mit der Iterations-Matrix $\mathbf{M} = \mathbf{I} - (2\eta/m)\mathbf{X}^\top\mathbf{X}$. Hier ist $\mathbf{I}$ die Einheits-Matrix, η die Lernrate, m die Anzahl der Datenpunkte und $\mathbf{X}$ die Merkmalsmatrix des Trainingsdatensatzs (siehe lineare Regression). Eine solche iteration konvergiert gegen den eindeutigen fixed point $\widehat{\mathbf{w}}$ von $\mathcal{F}^{(\eta)}$, für jede Initialisierung, genau dann, wenn der spectral radius ρ von $\mathbf{M}$, der größte Betrag eines Eigenwerts von $\mathbf{M}$, die Bedingung $\rho < 1$ erfüllt (Horn and Johnson, 2013, Thm. 5.6.12). Für eine diagonalizable $\mathbf{M}$ zeigt die Eigenwertzerlegung, warum: in der basis ihrer eigenvectors skaliert jede iteration die Komponenten des Fehlers $\mathbf{w}^{(t)} - \widehat{\mathbf{w}}$ mit den Eigenwerten, sodass der Fehler genau dann schrumpft, wenn jeder Eigenwert einen Betrag kleiner als eins hat.

Über die convergence iterativer Verfahren hinaus werden Eigenwerte auch verwendet, um die Konnektivität eines undirected graphen $\mathcal{G}$ mit der Anzahl der Knoten n zu messen. Die Laplacian matrix $\mathbf{L}^{(\mathcal{G})}$ von $\mathcal{G}$ ist positive semi-definite (psd), mit Eigenwerten $0 = \lambda_1 \leq \lambda_2 \leq \ldots \leq \lambda_n$ (von Luxburg, 2007, Prop. 1). Der zweitkleinste Eigenwert λ_2, die algebraic connectivity von $\mathcal{G}$, quantifiziert die Konnektivität: $\lambda_2 > 0$ genau dann, wenn $\mathcal{G}$ connected ist (von Luxburg, 2007, Prop. 2), und λ_2 nahe Null deutet auf zwei clusters von Knoten mit wenigen Kanten zwischen ihnen hin (Chung, 1997; Fiedler, 1973; von Luxburg, 2007, Sect. 5.1). Abb. 2 zeigt drei Kantensätze auf denselben sechs Knoten: das Hinzufügen von Kanten erhöht λ_2 von 0, für zwei clusters ohne verbindende Kante, auf 6, für den vollständigen graph. Der Wert 6 ist der größtmögliche. Für einen graph mit Einheits-edge weights addieren sich die Laplacian matrix von $\mathcal{G}$ und die des Komplements $\bar{\mathcal{G}}$, das die Kanten besitzt, die $\mathcal{G}$ fehlen, zu $n\mathbf{I} - \mathbf{1}\mathbf{1}^\top$. Ein Einheitsnorm-eigenvector $\mathbf{x}$ eines positiven Eigenwerts ist orthogonal zum konstanten vector $\mathbf{1}$, einem eigenvector des Eigenwerts 0, daher $\mathbf{x}^\top\mathbf{L}^{(\mathcal{G})}\mathbf{x} = n -$

$\mathbf{x}^\top \mathbf{L}^{(\bar{\mathcal{G}})} \mathbf{x} \leq n$, da $\mathbf{L}^{(\bar{\mathcal{G}})}$ psd ist. Für den vollständigen graph hat das Komplement keine Kanten, daher $\mathbf{L}^{(\mathcal{G})} = n\mathbf{I} - \mathbf{1}\mathbf{1}^\top$ und jeder Eigenwert außer $\lambda_1 = 0$ ist gleich n.



$\lambda_2 = 0$ (a) $\lambda_2 \approx 0.44$ (b) $\lambda_2 = 6$ (c)

Abbildung 2: Drei Kantensätze auf denselben sechs Knoten, mit zunehmender Konnektivität von links nach rechts, und dem zweitkleinsten Eigenwert λ_2 der entsprechenden Laplacian matrix. (a) Zwei clusters ohne verbindende Kante, sodass der graph unzusammenhängend ist und $\lambda_2 = 0$. (b) Eine Kante verbindet die clusters, sodass der graph connected ist und $\lambda_2 \approx 0.44$ positiv, aber nahe Null ist. (c) Der vollständige graph, bei dem jedes Knotenpaar verbunden ist, sodass der graph connected ist und $\lambda_2 = 6$ den größtmöglichen Wert, die Anzahl der Knoten $n = 6$, erreicht, da kein Eigenwert der Laplacian matrix eines graphen mit n Knoten und Einheits-edge weights n überschreitet. Daten generiert von `pythondemos/eigenvalue.py`

Der eigenvector, der λ_2 entspricht, als Fiedler vector bezeichnet, kann verwendet werden, um $\mathcal{G}$ in zwei clusters zu partitionieren: jeder Knoten wird entsprechend dem Vorzeichen seiner eigenen Komponente des Fiedler vectors zugewiesen (Fiedler, 1975; Spielman, 2025).

Siehe auch: Matrix, eigenvector, Eigenwertzerlegung, spectral radius, Gradientenabstieg, lineare Regression, contractive operator, fixed-point iteration, Laplacian matrix, algebraic connectivity, Fiedler vector, spectral clustering.

Querverweise

Literatur

1. Chung (1997). Spectral Graph Theory. American Mathematical Society, Providence, RI, USA.

2. Fiedler (1973). Algebraic connectivity of graphs. Czechoslovak Mathematical Journal, 23(2), pp. 298–305.

3. Fiedler (1975). A property of eigenvectors of nonnegative symmetric matrices and its application to graph theory. Czechoslovak Mathematical Journal, 25(4), pp. 619–633.

4. Horn and Johnson (2013). Matrix Analysis. 2nd ed., Cambridge Univ. Press, New York, NY, USA.

5. von Luxburg (2007). A tutorial on spectral clustering. Statist. Comput., 17(4), pp. 395–416.

6. Spielman (2025). Spectral and Algebraic Graph Theory (Incomplete Draft). Ebook.

7. Strang (2016). Introduction to Linear Algebra. 5th ed., Wellesley-Cambridge Press, Wellesley, MA, USA.