Eigenwerte und Singulärwertzerlegung

Magnus Vergin

21.01.2025

Gliederung

Eigenwerte und Singulärwertzerlegung

Ein Eigenwert und ein Eigenvektor einer quadratischen Matrix $\mathbf{A}$ sind ein Skalar λ und ein von Null verschiedener Vektor $\mathbf{x}$, sodass gilt:

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x} \tag{1}$$

Ein Singulärwert und ein Paar von Singulärvektoren einer quadratischen oder rechteckigen Matrix $\mathbf{A}$ sind ein nicht-negativer Skalar σ sowie zwei von Null verschiedene Vektoren $\mathbf{u}$ und $\mathbf{v}$, sodass gilt:

$$\mathbf{A}\mathbf{v} = \sigma\mathbf{u}, \quad \mathbf{A}^H\mathbf{u} = \sigma\mathbf{v} \tag{2}$$

$\mathbf{A}^H$ steht für die Hermitische Transposition, was die komplex-konjugierte Transposition einer komplexen Matrix bezeichnet

Eigenwertgleichung

Die Eigenwertgleichung für eine quadratische Matrix lässt sich schreiben als:

$$(\mathbf{A} - \lambda \mathbf{I})\mathbf{x} = 0, \quad \mathbf{x} \neq 0 \tag{3}$$

Dies impliziert, dass $\mathbf{A} - \lambda \mathbf{I}$ singulär ist und daher:

$$\det(\mathbf{A} - \lambda \mathbf{I}) = 0 \tag{4}$$

- Charakteristisches Polynom von $\mathbf{A}$
- Grad des Polynoms entspricht der Ordnung der Matrix
- n-$\times$-n-Matrix hat n Eigenwerte

Eigenwertzerlegung

- Seien $\lambda_1, \lambda_2, \ldots, \lambda_n$ die Eigenwerte einer Matrix $\mathbf{A}$
- $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ die zugehörigen Eigenvektoren.
- $\mathbf{\Lambda}$ als $n \times n$-Diagonalmatrix der Eigenwerte λ_j auf der Diagonale
- $\mathbf{X}$ als $n \times n$-Matrix, deren j-te Spalte $\mathbf{x}_j$ ist
- dann gilt:

$$\mathbf{AX} = \mathbf{X\Lambda}. \tag{5}$$

Eigenwertzerlegung

Wenn Eigenvektoren linear unabhängig sind und $\mathbf{X}^{-1}$ existiert, dann gilt:

$$\mathbf{A} = \mathbf{X}\boldsymbol{\Lambda}\mathbf{X}^{-1}. \tag{6}$$

Falls die Eigenvektoren von $\mathbf{A}$ nicht linear unabhängig sind, existiert eine solche Diagonale Zerlegung nicht.

Singulärwertzerlegung

Gleichung für Singulärwerte und -vektoren lassen sich in Matrixform schreiben als:

$$\mathbf{AV} = \mathbf{U\Sigma}, \quad \mathbf{A}^H\mathbf{U} = \mathbf{V\Sigma}^H \tag{7}$$

- $\mathbf{\Sigma}$: Diagonalmatrix mit derselben Dimensionen wie $\mathbf{A}$, Hauptdiagonale mit Singulärwerten, Rest Nullen
- $\mathbf{U}, \mathbf{V}$: Matrizen mit normierten Singulärvektoren als Spalten
- $\mathbf{U}^H\mathbf{U} = \mathbf{I}, \quad \mathbf{V}^H\mathbf{V} = \mathbf{I}$
- $\mathbf{U}, \mathbf{V}$ sind orthogonal (reell) oder unitär (komplex)

$$\mathbf{A} = \mathbf{U}\Sigma\mathbf{V}^H \tag{8}$$

Bedeutung von Eigen- und Singulärwerten

- Eigenwerte sind relevant, wenn eine $n \times n$-Matrix $\mathbf{A}$ als Abbildung eines n-dimensionalen Raums auf sich selbst betrachtet wird.
- Ziel: Basis finden, sodass die Matrix diagonal wird.
- Singulärwerte sind relevant, wenn eine $m \times n$-Matrix $\mathbf{A}$ eine Abbildung eines n-dimensionalen Raums in einen m-dimensionalen Raum darstellt.
- Ziel: Basis im Definitionsbereich finden, um eine möglicherweise andere Basis im Zielbereich zu definieren, sodass die Matrix diagonal wird.

Vollständige und ökonomische SVD

- Falls A eine $m \times n$-Matrix ist und $m > n$, dann ist U in der vollständigen SVD eine große, quadratische $m \times m$-Matrix. Die letzten $m - n$ Spalten von U sind überflüssig und werden nicht benötigt, um A zu rekonstruieren.
- Eine alternative Version der SVD, die Speicherplatz spart, ist die sogenannte **ökonomische SVD**.
    - In der ökonomischen Version werden nur die ersten n Spalten von U und die ersten n Zeilen von Σ berechnet.
    - Die Matrix V bleibt die gleiche $n \times n$-Matrix in beiden Zerlegungen.
- Beide Zerlegungen können als $A = U\Sigma V^H$ geschrieben werden, auch wenn U und Σ in der ökonomischen Zerlegung Untermatrizen der vollständigen Zerlegung sind.

Vollständige und ökonomische SVD

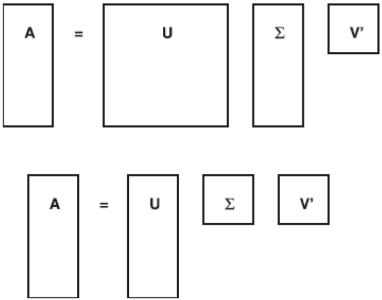

Figure 10.1. *Full and economy SVDs.*

Figure: 1

Beispiel

Die Matrix **A** lautet:

$$\mathbf{A} = \begin{bmatrix} -149 & -50 & -154 \\ 537 & 180 & 546 \\ -27 & -9 & -25 \end{bmatrix}$$

- Das charakteristische Polynom ist:

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \lambda^3 - 6\lambda^2 + 11\lambda - 6 = (\lambda - 1)(\lambda - 2)(\lambda - 3)$$

- Eigenwerte: $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$

Beispiel

$[X, \Lambda] = eig(A)$

$$\mathbf{\Lambda} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

$$\mathbf{X} = \begin{bmatrix} 0.3162 & -0.4041 & -0.1391 \\ -0.9487 & 0.9091 & 0.9740 \\ 0 & -0.2500 & 1.2857 \end{bmatrix}$$

$normalize(X, 1,' scale',' first')$

$$\mathbf{X} = \begin{bmatrix} 1 & 1 & 1 \\ -3 & -2.2500 & -7 \\ 0 & -0.2500 & 1.2857 \end{bmatrix}$$

Beispiel

$$\boldsymbol{\Lambda} = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 3 \end{bmatrix}$$

Durch Skalierung der zweiten und dritten Spalte:

$$\mathbf{X} = \begin{bmatrix} 1 & -4 & 7 \\ -3 & 9 & -49 \\ 0 & 1 & 9 \end{bmatrix}$$

$inv(X)$

$$\mathbf{X}^{-1} = \begin{bmatrix} 130 & 43 & 133 \\ 27 & 9 & 28 \\ -3 & -1 & -3 \end{bmatrix}$$

Diese Matrizen liefern die Eigenwertzerlegung unseres Beispiels:

$$\mathbf{A} = \mathbf{X}\boldsymbol{\Lambda}\mathbf{X}^{-\mathbf{1}}.$$

Beispiel

[U,S,V] = svd(A)

$$\mathbf{U} = \begin{bmatrix} -0.2691 & -0.6798 & 0.6822 \\ 0.9620 & -0.1557 & 0.2243 \\ -0.0463 & 0.7167 & 0.6959 \end{bmatrix}$$

$$\mathbf{S} = \begin{bmatrix} 817.7597 & 0 & 0 \\ 0 & 2.4750 & 0 \\ 0 & 0 & 0.0030 \end{bmatrix}$$

$$\mathbf{V} = \begin{bmatrix} 0.6823 & -0.6671 & 0.2990 \\ 0.2287 & -0.1937 & -0.9540 \\ 0.6944 & 0.7193 & 0.0204 \end{bmatrix}$$

Charakteristisches Polynom

Sei $\mathbf{A}$ eine 20×20-Diagonalmatrix mit $1, 2, \ldots, 20$ auf der Diagonale.

- Eigenwerte von $\mathbf{A}$ sind die Diagonalelemente.
- $\det(\mathbf{A} - \lambda \mathbf{I})$ ergibt:

$$
\begin{aligned}
P(\lambda) = \lambda^{20} &- 210\lambda^{19} + 20615\lambda^{18} - 1256850\lambda^{17} + 53327946\lambda^{16} \\
&- 1672280820\lambda^{15} + 40171771630\lambda^{14} - 756111184500\lambda^{13} \\
&+ 11310276995381\lambda^{12} - 135585182899530\lambda^{11} \\
&+ 1307535010540395\lambda^{10} - 10142299865511450\lambda^{9} \\
&+ 63030812099294896\lambda^{8} - 311333643161390640\lambda^{7} \\
&+ 1206647803780373360\lambda^{6} - 3599979517947607200\lambda^{5} \\
&+ 8037811822645051776\lambda^{4} - 12870931245150988800\lambda^{3} \\
&+ 13803759753640704000\lambda^{2} - 8752948036761600000\lambda \\
&+ 2432902008176640000.
\end{aligned}
$$

Charakteristisches Polynom

- Berechnung eines Polynoms in Potenzform ist anfällig für Rundungsfehler
- Speichern der Koeffizienten als IEEE-Gleitkommazahl verändert fünf Koeffizienten
- Beispielsweise ändern sich die letzten 3 Ziffern bei λ^4 von 776 zu 392

Charakteristisches Polynom

Auf 16 signifikante Stellen berechnete Koeffizenten:

1.00000000000000	10.99962843024064
2.00000000000096	12.00054374363591
2.99999999986640	12.99938073455790
4.00000000495944	14.00054798867380
4.99999991473414	14.99962658217055
6.00000084571661	16.00019208303847
6.99999455544845	16.99992773461773
8.00002443256894	18.00001875170604
8.99992001186835	18.99999699774389
10.00019696490537	20.00000022354640

Symmetrische und Hermitesche Matrizen

- Eine reelle Matrix ist symmetrisch, wenn sie gleich ihrer Transponierten ist:

$$\mathbf{A} = \mathbf{A}^T \tag{9}$$

- Eine komplexe Matrix ist Hermitesch, wenn sie gleich ihrer komplex-konjugierten Transponierten ist:

$$\mathbf{A} = \mathbf{A}^H \tag{10}$$

Eigenwerte und Eigenvektoren

- Die Eigenwerte und Eigenvektoren einer reellen symmetrischen Matrix sind reell.
- Die Matrix der Eigenvektoren kann so gewählt werden, dass sie orthogonal ist. Daher lautet die Eigenwertzerlegung für A, wenn $A = A^T$:

$$\mathbf{A} = \mathbf{X}\Lambda\mathbf{X}^T \tag{11}$$

wobei $X^T X = I$ und Λ eine Diagonalmatrix ist.

Eigenwerte und Eigenvektoren

- Die Eigenwerte einer Hermiteschen Matrix sind reell, auch wenn die Eigenvektoren komplex sein können.
- Die Matrix der Eigenvektoren kann so gewählt werden, dass sie unitär ist. Daher lautet die Eigenwertzerlegung für A, wenn $A = A^H$:

$$\mathbf{A} = \mathbf{X}\Lambda\mathbf{X}^H \tag{12}$$

wobei $X^H X = I$ und Λ reell ist.

Beziehung zwischen Eigen- und Singulärwerten

- Für symmetrische und hermitesche Matrizen sind die Eigenwerte und Singulärwerte eng miteinander verknüpft.
- Ein nicht-negativer Eigenwert ($\lambda \geq 0$) ist auch ein Singulärwert:

$$\sigma = \lambda \tag{13}$$

- Die zugehörigen Vektoren sind:

$$\mathbf{u} = \mathbf{v} = \mathbf{x} \tag{14}$$

- Ein negativer Eigenwert ($\lambda < 0$) wird in der Singulärwertzerlegung als dessen Absolutwert dargestellt:

$$\sigma = |\lambda| \tag{15}$$

- Ein zugehöriger Singulärvektor ist das Negative des anderen:

$$\mathbf{u} = -\mathbf{v} = \mathbf{x} \tag{16}$$

Empfindlichkeit und Genauigkeit von Eigenwerten

- Eigenwerte einer Matrix sind empfindlich gegenüber Störungen.
- Angenommen, A besitzt eine vollständige Menge linear unabhängiger Eigenvektoren, sodass die Eigenwertzerlegung

$$A = X \Lambda X^{-1} \tag{17}$$

  existiert. Umgeschrieben:

$$\Lambda = X^{-1} A X \tag{18}$$

- Nun sei δA eine kleine Änderung in A (z. B. durch Rundungsfehler). Dann ergibt sich:

$$\Lambda + \delta\Lambda = X^{-1}(A + \delta A)X \tag{19}$$

- Daraus folgt:

$$\delta\Lambda = X^{-1} \delta A X \tag{20}$$

Empfindlichkeit und Genauigkeit von Eigenwerte

- Nimmt man die Norm, erhält man:

$$|\delta\Lambda\| \leq \|X^{-1}\|\|X\|\|\delta A\| = \kappa(X)\|\delta A\| \tag{21}$$

- Dabei ist $\kappa(X) = \|X^{-1}\|\|X\|$ die Konditionszahl der Eigenvektormatrix.
- Dies führt zu der Schlussfolgerung:

  Die Empfindlichkeit der Eigenwerte wird durch die Konditionszahl der Eigenvektormatrix $\kappa(X)$ bestimmt, nicht durch die Kondition der Matrix A selbst.

Beispiel

- In MATLAB kann man den Befehl `condest` zur Schätzung der Konditionszahl der Eigenvektormatrix nutzen.
- Für die Eigenvektormatrix $\mathbf{X}$ der Matrix

$$\mathbf{A} = \begin{bmatrix} -149 & -50 & -154 \\ 537 & 180 & 546 \\ -27 & -9 & -25 \end{bmatrix}$$

  gibt der Befehl `condest(X)` das Ergebnis 1.2002×10^3.
- Eine Störung der Matrix A könnte Störungen der Eigenwerte resultieren, die 1.2×10^3 mal so groß sind.
- Die Eigenwerte der Matrix A sind schlecht konditioniert.

Empfindlichkeit eines Eigenwerts und Eigenwert-Konditionszahl

- Durch detaillierte Analyse kann die Empfindlichkeit eines einzelnen Eigenwerts betrachtet werden.
- Die Eigenwert-Konditionszahl für einen spezifischen Eigenwert λ ist definiert als:

$$\kappa(\lambda, A) = \frac{\|y\|\|x\|}{|y^H x|} \tag{22}$$

- Dabei gilt:
  - y: Linker Eigenvektor zu λ,
  - x: Rechter Eigenvektor zu λ,
  - $y^H x$: Das Skalarprodukt der beiden Vektoren.

Empfindlichkeit eines Eigenwerts und Eigenwert-Konditionszahl

- Wenn $y^H x = 1$ (durch Normierung), dann lautet die Konditionszahl:

$$\kappa(\lambda, A) = \|y\|\|x\| \tag{23}$$

- Da $x \leq X$ und $y \leq X^{-1}$, gilt:

$$\kappa(\lambda, A) \leq \kappa(X) \tag{24}$$

Konditionszahl des Eigenwerts in MATLAB

- In MATLAB kann man mit dem Befehl `condeig` die Konditionszahl des Eigenwerts berechnen.
- Wenn wir bei unserem Beispiel der Matrix **A** bleiben, ergibt das für die Eigenwerte $\lambda = [1.0000, 2.0000, 3.0000]$:

$$\kappa = [603.6390, 395.2366, 219.2920]$$

- Die Ergebnisse lassen darauf schließen, dass $\lambda_1 = 1$ etwas fehleranfälliger ist als $\lambda_2 = 2$ und $\lambda_3 = 3$.

Eigenschaften symmetrischer und hermitescher Matrizen

- Wenn A reell und symmetrisch, oder komplex und hermitesch ist, dann sind der linke und rechte Eigenvektor derselbe.
- Also gilt für symmetrische und hermitesche Matrizen:

$$\kappa(\lambda, A) = 1 \tag{25}$$

- Eigenwerte symmetrischer und hermitescher Matrizen sind wohl konditioniert.
- Störungen der Matrix führen zu Störungen der Eigenwerte in der gleichen Größenordnung.

Mehrfache Eigenwerte und ihre Auswirkungen

- Wenn λ_k ein mehrfacher Eigenwert ist,müssen die Eigenvektoren nicht linear unabhängig sein.
- Das charakteristische Polynom $p(\lambda)$ lässt sich schreiben als:

$$p(\lambda) = \det(A - \lambda I) = (\lambda - \lambda_k)^m q(\lambda) \tag{26}$$

- Dabei ist m die Multiplizität von λ_k und $q(\lambda)$ ein Polynom vom Grad $n - m$.

Mehrfache Eigenwerte und ihre Auswirkungen

- Eine Störung der Matrix der Größenordnung δ führt zu einer Änderung des charakteristischen Polynoms $p(\lambda) = 0$ zu:

$$p(\lambda) = O(\delta) \tag{27}$$

- Daher gilt:

$$(\lambda - \lambda_k)^m = \frac{q(\lambda)}{O(\delta)} \tag{28}$$

- Die Eigenwerte ändern sich wie folgt:

$$\lambda = \lambda_k + O(\delta^{1/m}) \tag{29}$$

Beispiel

- Die Matrix A sieht folgendermaßen aus:

$$A = \begin{pmatrix} 2 & 1 & 0 & 0 & \cdots & 0 \\ 0 & 2 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 2 & 1 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & 1 \\ \delta & 0 & 0 & 0 & \cdots & 2 \end{pmatrix}$$

- Das charakteristische Polynom lautet:

$$(\lambda - 2)^{16} = \delta \tag{30}$$

Beispiel

- Wenn $\delta = 0$, dann gibt es einen Eigenwert der Multiplizität 16 bei $\lambda = 2$.
- Wenn δ der Größenordnung eines Rundungsfehlers entspricht, also $\delta \approx 10^{-16}$, dann liegen die Eigenwerte auf einem Kreis in der komplexen Ebene:

$$\text{Mittelpunkt} = 2, \quad \text{Radius} = (10^{-16})^{1/16} = 0.1$$

- Eine Störung durch einen Rundungsfehler ändert die Eigenwerte auf 16 verschiedene Werte, z.B.:

$$\lambda = 1.9,\ 2.1,\ 2.0924 + 0.0383i$$

Empfindlichkeit und Genauigkeit von Singulärwerte

- Die Empfindlichkeit von Singulärwerten ist wesentlich einfacher zu charakterisieren.
- Das Singulärwertproblem ist immer wohl konditioniert.
- Die Störungsanalyse führt zu:

$$\Sigma + \delta\Sigma = U^H(A + \delta A)V \tag{31}$$

- Da U und V orthogonal oder unitär sind, gilt:

$$\|\delta\Sigma\| = \|\delta A\| \tag{32}$$

- Störungen beliebiger Größe in einer Matrix bewirken Störungen derselben Größe in ihren Singulärwerten.

Empfindlichkeit und Genauigkeit von Singulärwerte

- Störungen und Genauigkeit werden relativ zur Norm der Matrix gemessen, die durch den größten Singulärwert σ_1 gegeben ist:

$$\|A\|_2 = \sigma_1 \tag{33}$$

- Wenn die Singulärwerte sich über mehrere Größenordnungen erstrecken, können die kleineren Singulärwerte nicht die volle relative Genauigkeit erreichen.
- Insbesondere gilt:
  - Wenn die Matrix singulär ist, sind einige der σ_i genau null.
  - Die berechneten Werte dieser σ_i liegen typischerweise in der Größenordnung von $\varepsilon\|A\|$, wobei ε der Gleitkomma-Genauigkeitsparameter ist.

QR-Algorithmus

- Der QR-Algorithmus basiert auf der wiederholten Anwendung der QR-Faktorisierung.
- Jede Matrix A kann in das Produkt einer orthogonalen Matrix Q und einer oberen Dreiecksmatrix R zerlegt werden:

$$A = QR \tag{34}$$

QR-Algorithmus - Beispiel

n = size(A,1), I = eye(n,n)
s = A(n,n); [Q,R] = qr(A - s*I); A = R*Q + s*I

$$\mathbf{A} = \begin{bmatrix} -149 & -50 & -154 \\ 537 & 180 & 546 \\ -27 & -9 & -25 \end{bmatrix}$$

Nach der ersten Iteration:

$$\begin{pmatrix} 28.8263 & -259.8671 & 773.9292 \\ 1.0353 & -8.6686 & 33.1759 \\ -0.5973 & 5.5786 & -14.1578 \end{pmatrix}$$

QR-Algorithmus - Beispiel

Nach fünf weiteren Iterationen:

$$\begin{pmatrix} 2.7137 & -10.5427 & -814.0932 \\ -0.0767 & 1.4719 & -76.5847 \\ 0.0006 & -0.0039 & 1.8144 \end{pmatrix}$$

Nach fünf weiteren Iterationen:

$$\begin{pmatrix} 3.0716 & -7.6952 & 802.1201 \\ 0.0193 & 0.9284 & 158.9556 \\ 0.0000 & 0.0000 & 2.0000 \end{pmatrix}$$

Optimierungen des QR-Algorithmus

- Der QR-Algorithmus wird in der Praxis nicht in dieser einfachen Form verwendet. Er enthält folgende Optimierungen:
- **Reduktion zur Hessenberg-Form**:
  - Bevor der Algorithmus beginnt, wird A auf Hessenberg-Form reduziert, wobei alle Elemente unterhalb der ersten Subdiagonale null sind.
  - Dies beschleunigt die Berechnung erheblich.

Verbesserte Shift-Strategien im QR-Algorithmus

- Moderne Varianten des QR-Algorithmus verwenden weiterentwickelte Shift-Strategien, die je nach Typ der Matrix variieren:
- **Symmetrische Matrizen:**
  - Eine tridiagonale Reduktion ermöglicht eine effiziente Berechnung
  - QR-Algorithnus konvergiert immer
  - keine Beispiele bekannt, die trotz Rundungsfehlern dazu führen, dass Implimentierung in Matlab fehlschlägt

SVD-Variante des QR-Algorithmus

- Für die Berechnung der Singulärwerte wird der Algorithmus so angepasst, dass er eine bidiagonale Form erhält, die die Singulärwerte bewahrt.
- dieselbe garantierte Konvergenzeigenschaft wie die symmetrische Eigenwertiteration

Literatur

- Moler, C. B., Numerical Computing with Matlab, SIAM, 2004.
- Voß, H., Grundlagen der numerischen Mathematik, Technische Universität hamburg-Harburg, 2004
- Taboga, Marco (2021). "Linear independence of eigenvectors", Lectures on matrix algebra. https://www.statlect.com/matrix-algebra/linear-independence-of-eigenvectors.