

Fehlen noch Quellen natürlich Dynamic Mode Decomposition, kurz DMD, ist ein datengetriebener Algorithmus, der ursprünglich im Feld der Strömungsdynamik entwickelt wurde. Grundsätzlich dient er der Analyse dynamischer Systeme mittels Singulärwertzerlegung von großen Datenmatrizen. DMD kann anhand von Messdaten mehrerer Zeitpunkte die Dynamik eines Systems approximieren. Die Genauigkeit hängt unter anderem davon ab, wie viele Informationen in Form von Messdaten über das System zur Verfügung stehen und wie groß die Messfehler also das Rauschen der Daten ist. Dabei ist DMD gut im Erkennen von Mustern, wie im Falle der Strömungsdynamik z.B. Verwirbelungen. Probleme treten jedoch auf, wenn man DMD außerhalb von Simulationen verwendet und echte Messungen analysieren will. Bei Messungen hat man immer ein Rauschen in den Daten, welche das Ergebnis verfälschen. Dadurch kommt es durch Fehlerfortpflanzung schnell zu größeren Ungenauigkeiten. Im Laufe der letzten Jahre wurden eine Vielzahl von Varianten der DMD Algorithmen entwickelt, um unterschiedliche Schwächen des Algorithmus auszugleichen. Das klassische DMD benötigt Messdaten in einheitlichen Zeitabständen. Für jede Messung werden alle Daten in eine Spalte x_i einer Datenmatrix X eingetragen. Zusätzlich legen wir eine zweite Datenmatrix X' so an, dass wir in X die Messungen (x_1, \dots, x_{k-1}) und in X' die Messungen (x_2, \dots, x_k) haben. Die Datenmatrizen sind in der Regel $n \times m$ große Matrizen, wobei sie deutlich mehr Zeilen als Spalten haben, also $n \gg m$ gilt. Das Ziel von DMD ist es nun einen linearen Operator A bzw. seine Spektralzerlegung zu finden, welcher das System

$$X' \approx AX$$

so gut wie möglich beschreibt. Der dafür am besten passende Operator A lässt sich durch

$$A = \arg \min_A \|X' - AX\|_F = X'X^\dagger$$

bestimmen. Dabei sind $\|\cdot\|_F$ die Frobeniusnorm und X^\dagger die Moore-Penrose-Inverse, welche wir später noch genauer beschreiben werden. Die Moore-Penrose-Inverse lässt sich dann mittels Singulärwertzerlegung (SWZ) der Datenmatrix X bestimmen. Die SWZ beschreiben wir auch später nochmal genauer, liefern aber für einen kleinen Überblick eine vereinfachte Beschreibung. Bei der SWZ finden wir für ein $X \in \mathbb{R}^{n \times m}$ zwei orthogonale Matrizen U und V mit passenden Dimensionen und $\hat{\Sigma} = \text{diag}(\sigma_1, \dots, \sigma_r)$ mit

$r \leq \min(n, m)$ und $\sigma_1 \geq \sigma_2 \geq \dots \geq \sigma_r > 0$ so, dass

$$X = U\Sigma V^\top \text{ mit } \Sigma = \begin{pmatrix} \hat{\Sigma} & 0 \\ 0 & 0 \end{pmatrix} \in \mathbb{R}^{n \times m}$$

gilt. Bei der Bestimmung von A wird der Rechenaufwand schnell zu groß, weshalb der DMD Algorithmus eine Dimensionsreduktion durchführt. Damit kann man eine Spektralzerlegung von A machen, ohne A direkt zu benutzen. Durch die oben genannte Form der Datenmatrizen kann man direkt sehen, dass X maximal von Rang m sein kann. Es gibt also maximal $r \leq m$ Singulärwerte, die nicht Null sind. Reduzieren wir also unser Σ aus der SWZ auf eine $r \times r$ große Matrix, und entfernen bis auf die ersten r Spalten alle Spalten von U und V , erhalten wir mit den reduzierten Matrizen U_r , Σ_r und V_r ,

$$X_r = U_r \Sigma_r V_r^\top = U_r \hat{\Sigma} V_r^\top .$$

Nach dem Eckart-Young-Mirsky-Theorem **Quelle!! Vielleicht doch eher MSE nehmen weil gleich wie im paper** lässt sich durch

$$\|X_r\|_F^2 = \sum_{i=1}^r \sigma_i^2$$

zeigen, dass X_r die beste Rang r -Approximation bezüglich der Frobeniusnorm ist. Die Spektralzerlegung von A können wir damit über die viel kleinere Matrix

$$\tilde{A} = U_r^\top A U_r$$

bestimmen. Diese Methode der Dimensionsreduktion für SWZ, ist als ökonomischer SWZ bekannt.

Eine weitere Methode, welche im Paper [1] von Matan Gavish und David L. Donoho behandelt wird, ist das „singular value hard thresholding“ (SVHD). Das ist eine Rangreduzierung bis alle Singulärwerte über einem definierten Schwellenwert τ liegen. Gavish und Donoho definieren einer Rahmenstruktur, in der die Datenmatrix X durch

$$X := X^{\text{echt}} + \gamma X^{\text{rauschen}}$$

beschrieben wird. Wir haben also X in einen echten Teil und einem durch γ skalierten Fehleranteil aufgeteilt.

nicht sicher wie detailliert ich die Rahmenstruktur definieren soll oder ob überhaupt.

Innerhalb dieser Rahmenstruktur haben Gavish und Donoho herausgefunden, dass für bekanntes Rauschen γ der optimale Schwellenwert τ_* für Singulärwerte bei

$$\tau_* = \lambda_*(y)\gamma\sqrt{n} \quad \text{mit} \quad \lambda_*(y) := \sqrt{2(y+1) + \frac{8y}{(y+1) + \sqrt{y^2 + 14y + 1}}}$$

liegt. Bei quadratischen Matrizen reduziert sich λ_* sogar auf $\frac{4}{\sqrt{3}}$.

In echten Messungen kommt es aber eher selten vor, dass das Rauschen bekannt ist.

Dafür kommt an dieser Stelle die Marčenko-Pastur-Verteilung zum einsatz. lalala

Plan: (notation aus paper bis auf $\sigma \rightarrow \gamma$)

1. Rahmenstruktur iid und co [✓]
2. für noise bekannt und $n \times n$ $\tau_* = \frac{4}{\sqrt{3}}\gamma\sqrt{n}$ [✓]
3. noise unbekannt und $n \times m$ oder andersrum maaaaaaaan keine ahnung gerade
 $\rightarrow \hat{\tau}_*(\beta, Y) := \lambda_*(\beta)\sqrt{n}\hat{\gamma}(Y) = \frac{\lambda_*(\beta)}{\sqrt{\mu_\beta}}y_{med}$
4. mit $\omega(\beta) = \frac{\lambda_*(\beta)}{\sqrt{\mu_\beta}}$ dann $\hat{\tau}_*(\beta, Y) = \omega(\beta)y_{med}$
5. dann μ_β aus MP
6. warum MP
7. was kann MP
8. ich mach MP