Proceeding Series of the Brazilian Society of Computational and Applied Mathematics, v. 7, n. 1, 2020.

Trabalho apresentado no XXXIX CNMAC, Uberlândia - MG, 2019.

Proceeding Series of the Brazilian Society of Computational and Applied Mathematics

Análise de Dados de Alta Dimensão Via Decomposição em Modos Dinâmicos

Lucy T. Takahashi [1]
Departamento de Matemática, UFJF, Minas Gerais, Brasil
Luis A. D'Afonseca [2]
Departamento de Matemática, CEFET, Minas Gerais, Brasil
Wilson C. Ferreira Jr. [3]
IMECC, UNICAMP, São Paulo, Brasil

1 Introdução

Neste trabalho o objetivo é apresentar uma aplicação da Decomposição em Modos Dinâmicos (DMD), além de sua estrutura e algoritmo [3]. O DMD é um método que foca na descoberta de modos espaço-temporais em dados de alta dimensão coletados de sistemas complexos com dinâmica de tempo. O algoritmo conecta: análise de sistemas não-lineares, dinâmica livre de equações e a capacidade de lidar eficientemente com dados de alta dimensão.

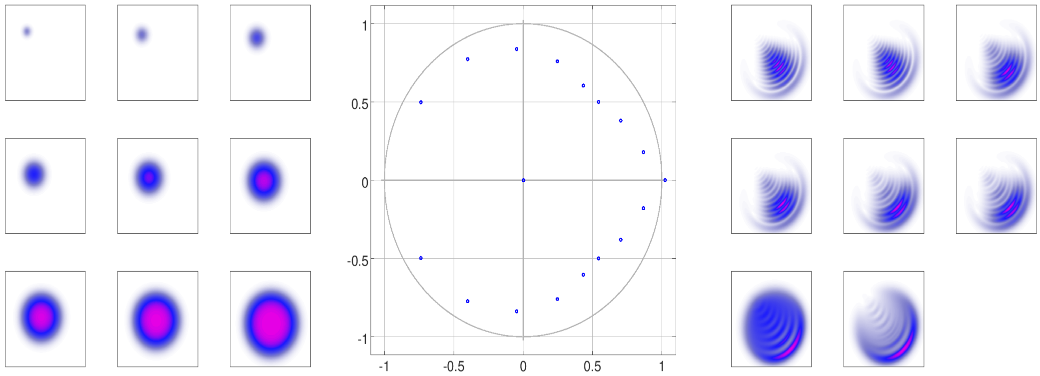
2 Decomposição em Modos Dinâmicos

A Decomposição em Modos Dinâmicos (DMD) é um método que analisa a relação entre sequências de medidas, no caso espaço-tempo. É fundamentalmente livre de equações operando apenas sobre *snapshots* no tempo. Os dados de entrada podem ser gerados a partir de simulações, experimentos ou dados históricos, logo oriundos de sistemas complexos e não-lineares. Essa análise consiste em realizar uma Análise dos Componentes Principais (PCA) para a redução da alta dimensão. Considerando $x \in \mathbb{R}^n$ e $k \in \mathbb{N}$, temos x_{k+1} os dados espaço-temporais de uma doença em uma medida futura e x_k numa medida anterior, chamados *snapshots*. Assumimos $x_{k+1} = Ax_k$ onde o operador $A \in \mathbb{R}^{n\times n}$ é o melhor ajuste para todos os pares de *snapshots* no sentido de quadrados mínimos. No caso de doenças $x_k = (x_k^1, x_k^2, \ldots, x_k^n)$ tem em cada entrada o número de casos, numa localidade $i = 1, \ldots, n$, no tempo k.

A relação entre os pares x_k e x_{k+1} pode ser descrita aproximadamente pela forma matricial: $X' \approx A\ X$. Daí, $A = X'X^\dagger = X'V\Sigma^{-1}U^* \approx \bar{A} = X'\widetilde{V}\widetilde{\Sigma}^{-1}\widetilde{U}^*$, tal que $\widetilde{\Sigma} \in \mathbb{R}^{r\times r}$ onde r é uma escolha para o truncamento. As características da dinâmica podem ser

[1] ltiemi@gmail.com
[2] luis.dafonseca@cefetmg.br
[3] wilson@ime.unicamp.br

encontradas pela autodecomposição: $\widetilde{A}W = W\Lambda$. A matriz $\widetilde{A}$ é obtida truncando a matriz A em r valores singulares, W contém seus respectivos autovetores e Λ os autovalores. Os modos dinâmicos que descrevem a dinâmica são construídos como: $\phi = X'\widetilde{V}\widetilde{\Sigma}^{-1}w$. Essa técnica foi utilizada em [2] para se determinar o estado atual da atividade da gripe nos USA, utilizaram dados agregados de pesquisa do *Google* e dados históricos da gripe.



(a) Dados modelados (b) Autovalores (c) Modos Dinâmicos

Figura 1: Análise dos componentes principais e reconstrução de uma propagação gerada numericamente.

Afim de aplicar o DMD geramos numericamente uma propagação espacial, modelamos a propagação de uma população que em cada célula segue o modelo de Verhulst e pode se mover para as células vizinhas. O *grid* foi formado com 100×100 células, como ilustrado na Figura 1a. Na Figura 1b exibimos os autovalores no plano complexo. Com base nos módulos dos autovalores determinamos os modos dinâmicos com maior e menor influência na propagação, exibidos na Figura 1c. Utilizando um truncamento de 95% de ajuste, observamos que o resultado é análogo aos dados reais. Assim, temos o DMD como uma ferramenta eficiente, sem equações, para descrever dados de alta dimensão e não lineares sobre a propagação de doenças.

Agradecimentos

À Capes pelo apoio financeiro.

Referências

[1] J. N. Kutz, *Data-Driven Modeling & Scientific Computation: Methods for Complex Systems & Big Data*, Oxford University Press, Oxford, 2013.

[2] J. N. Kutz, S. L. Bruton, B. W. Bruton and J. L. Proctor, *Dynamic Mode Decomposition: Data-Driven Moldeling of Complex Systems*, SIAM, Philadelphia, 2016.

[3] J. L. Proctor e P. A. Eckhoff. Discovering dynamic patterns from infectious disease data using dynamic mode decomposition, *Int. Health*, 7: 139-145, 2015. DOI: 10.1093/inthealth/ihv009.

