Tekil Değer Ayrıştırması

T.C. Trakya Üniversitesi
Mühendislik Fakültesi
Elektrik - Elektronik Mühendisliği Bölümü
Kontrol Anabilim Dalı

Dr. Öğr. Üyesi Işık İlber Sırmatel
sirmatel.github.io

Kaynak (source)

Lecture slides for Introduction to Linear Dynamical Systems

Stephen Boyd

Konu listesi

1. Tekil değer ayrıştırması

2. Tekil değer ayrıştırmasının yorumları

3. Tekil değer ayrıştırması uygulamaları

Bölüm 1

Tekil değer ayrıştırması

Tekil değer ayrıştırması (SVD)

bir $A \in \mathbb{R}^{m \times n}$ matrisinin **tekil değer ayrıştırması** (*singular value decomposition*) (SVD)

$$A = U\Sigma V^T$$

şeklinde ifade edilir

- bir matrisin tekil değer ayrıştırması daima mevcuttur
- $\mathrm{rank}(A) = r$
- $U \in \mathbb{R}^{m \times r}$, $U^T U = I$
- $V \in \mathbb{R}^{n \times r}$, $V^T V = I$
- $\Sigma = \mathrm{diag}(\sigma_1, \ldots, \sigma_r)$ $(\sigma_1 \geq \cdots \geq \sigma_r > 0)$

Tekil değer ayrıştırması (SVD)

U ve V'yi

$$U = \begin{bmatrix} u_1 & \ldots & u_r \end{bmatrix} \qquad V = \begin{bmatrix} v_1 & \ldots & v_r \end{bmatrix}$$

şeklinde (yani, sütunlarını belirterek) yazarsak

$$A = U\Sigma V^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

yazabiliriz

- σ_i, A'nın (sıfıra eşit olmayan) **tekil değer**leri (*singular value*)
- u_i, A'nın **sol tekil** (*left singular*) vektörleri
- v_i, A'nın **sağ tekil** (*right singular*) vektörleri

Tekil değer ayrıştırması (SVD)

$$A^TA = (U\Sigma V^T)^T(U\Sigma V^T) = V\Sigma^2 V^T$$

yazabiliriz

dolayısıyla:

- v_i, A^TA'nın (sıfıra eşit olmayan özdeğerlerine karşılık gelen) özvektörleridir
- $\sigma_i = \sqrt{\lambda_i(A^TA)}$ (ve, $i > r$ için $\lambda_i(A^TA) = 0$)
- $\|A\| = \sigma_1$ (yazıyla: bir matrisin (spektral) normu, o matrisin maksimum tekil değerine eşittir)

Tekil değer ayrıştırması (SVD)

benzer şekilde,

$$AA^T = (U\Sigma V^T)(U\Sigma V^T)^T = U\Sigma^2 U^T$$

yazabiliriz

dolayısıyla:

- u_i, AA^T'nın (sıfıra eşit olmayan özdeğerlerine karşılık gelen) özvektörleridir
- $\sigma_i = \sqrt{\lambda_i(AA^T)}$ (ve, $i > r$ için $\lambda_i(AA^T) = 0$)
- $u_1, \ldots, u_r$, $\mathcal{R}(A)$ (A'nın sütun uzayı) için bir birim dikgen taban oluşturur
- $v_1, \ldots, v_r$, $\mathcal{R}(A^T)$ (A'nın satır uzayı) için bir birim dikgen taban oluşturur

Ek bilgi: Tam tekil değer ayrıştırması

$A \in \mathbb{R}^{m\times n}$ $(\mathrm{rank}(A) = r)$ için

$$A = U_1\Sigma_1V_1^T = \begin{bmatrix} u_1 & \cdots & u_r \end{bmatrix} \begin{bmatrix} \sigma_1 & & \\ & \ddots & \\ & & \sigma_r \end{bmatrix} \begin{bmatrix} v_1^T \\ \vdots \\ v_r^T \end{bmatrix}$$

şeklindeki SVD'ye ekonomik (veya, **ince** (*thin*)) SVD denir

- $U = \begin{bmatrix} U_1 & U_2 \end{bmatrix} \in \mathbb{R}^{m\times m}$ ve $V = \begin{bmatrix} V_1 & V_2 \end{bmatrix} \in \mathbb{R}^{n\times n}$ olacak şekilde dikgen $U_2 \in \mathbb{R}^{m\times(m-r)}$ ve $V_2 \in \mathbb{R}^{n\times(n-r)}$ matrislerini bulalım
- Σ_1'e sıfır satır ve sütunlar ekleyerek $\Sigma \in \mathbb{R}^{m\times n}$'yı oluşturalım:

$$\Sigma = \left[\begin{array}{c|c} \Sigma_1 & 0_{r\times(n-r)} \\ \hline 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{array}\right]$$

Ek bilgi: Tam tekil değer ayrıştırması

buradan, A'nın **tam** (*full*) SVD'sini

$$A = U_1 \Sigma_1 V_1^T = \underbrace{\left[\begin{array}{c|c} U_1 & U_2 \end{array}\right]}_{U} \underbrace{\left[\begin{array}{c|c} \Sigma_1 & 0_{r\times(n-r)} \\ \hline 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{array}\right]}_{\Sigma} \underbrace{\left[\begin{array}{c} V_1^T \\ \hline V_2^T \end{array}\right]}_{V^T}$$

şeklinde yazabiliriz

Bölüm 2

Tekil değer ayrıştırmasının yorumları

SVD'nin eşleme adımları ile yorumu

$$A = U\Sigma V^T = \sum_{i=1}^{r} \sigma_i u_i v_i^T$$

$$\xrightarrow{x} \boxed{V^T} \xrightarrow{V^T x} \boxed{\Sigma} \xrightarrow{\Sigma V^T x} \boxed{U} \xrightarrow{Ax}$$

doğrusal eşleme (*mapping*) $y = Ax$

- x'in giriş yönleri $(v_1, \ldots, v_r)$ doğrultusundaki katsayılarını hesapla
- katsayıları σ_i ile **ölçekle** (*scale*)
- sonucu çıkış yönleri $(u_1, \ldots, u_r)$ doğrultusunda **yeniden oluştur** (*reconstitute*)

şeklinde adımlarına ayrıştırılabilir

SVD'nin eşleme adımları ile yorumu

not: simetrik A için SVD ile özayrışma arasındaki fark:
SVD'de giriş ve çıkış yönleri **farklı**

- v_1 en yüksek kazançlı giriş yönü
- u_1 en yüksek kazançlı çıkış yönü
- $Av_1 = \sigma_1 u_1$

SVD'nin eşleme adımları ile yorumu

SVD, giriş ve çıkış yönlerinin fonksiyonu olarak bir matrisin kazancıyla ilgili net bir fikir edinmemizi sağlar

örnek: $A \in \mathbb{R}^{4\times 4}$, $\Sigma = \mathrm{diag}(10,\ 7,\ 0.1,\ 0.05)$

- v_1 ve v_2 yönlerindeki giriş bileşenleri, A ile çarpımla kuvvetlendirilir (yaklaşık 10 kat) ve daha çok u_1 ve u_2'nin gerdiği düzlem boyunca çıkışa yansır
- v_3 ve v_4 yönlerindeki giriş bileşenleri, A ile çarpımla zayıflatılır (yaklaşık 10 kat)
- $\frac{\|Ax\|}{\|x\|}$'in değeri 10 ile 0.05 arasında değişir
- A tersi alınabilirdir
- bazı uygulamalar için "A'nın kertesi esasında 2" diyebiliriz

SVD'nin eşleme adımları ile yorumu

örnek: $A \in \mathbb{R}^{2\times 2}$, $\sigma_1 = 1$, $\sigma_2 = 0.5$

- x'i v_1 ve v_2 doğrultusunda çöz (yani, x'i v_1-v_2 koordinatlarında yaz):

$$v_1^T x = 0.6, \quad v_2^T x = 0.6 \longrightarrow x = 0.5v_1 + 0.6v_2$$

- sonucu $\sigma_1 = 1$ ve $\sigma_2 = 0.5$ ile ölçekle:

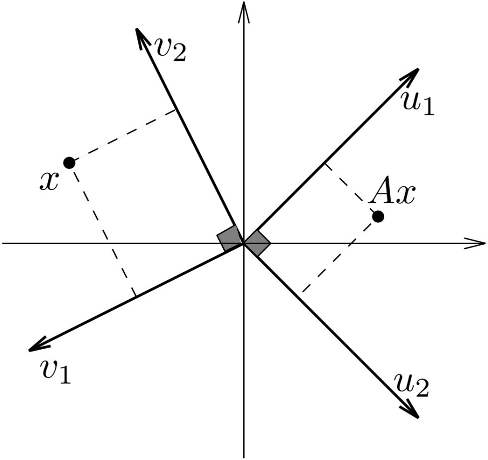
$$v_1^T x \sigma_1 = 0.6 \qquad v_2^T x \sigma_2 = 0.3$$

- u_1 ve u_2 doğrultusunda Ax'i oluştur:

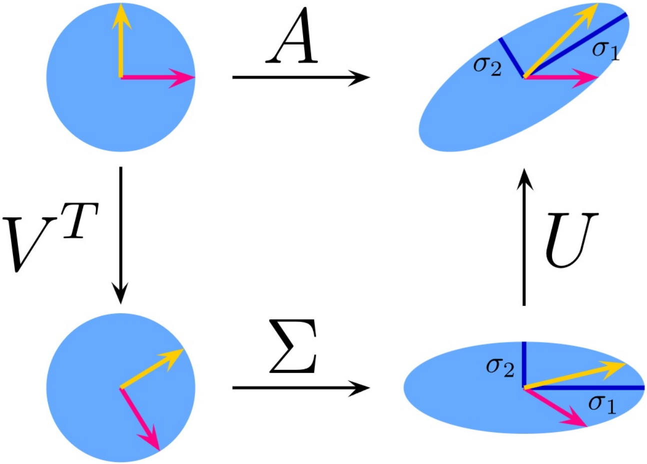
$$Ax = v_1^T x \sigma_1 u_1 + v_2^T x \sigma_2 u_2 = 0.5u_1 + 0.3u_2$$

SVD'nin eşleme adımları ile yorumu

örnek (devam):

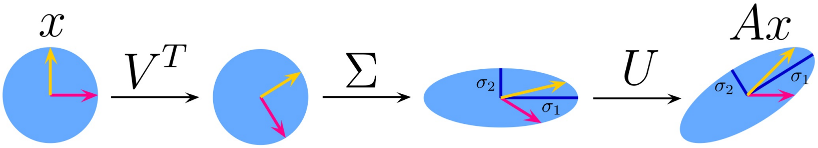


SVD'nin geometrik yorumu



$$A = U\Sigma V^T$$

SVD'nin geometrik yorumu



kaynak (source): Georg-Johann, CC BY-SA 3.0

- x'i V^T ile döndür
- sonucu Σ ile ölçekle
- sonucu U ile döndürerek Ax'i hesapla

Bölüm 3

Tekil değer ayrıştırması uygulamaları

Genel sözde ters

$A \neq 0$ olsun. A'nın SVD'si $A = U\Sigma V^T$ ise, A'nın **sözde ters**i (*pseudo-inverse*)

$$A^\dagger = V\Sigma^{-1}U^T$$

şeklinde yazılabilir

A uzun matris; tam sütun kerteli (yani, $(A^TA)^{-1}$ mevcut) ise

$$A^\dagger = (A^TA)^{-1}A^T$$

ile $Ax = b$'nin en küçük kareler yaklaşık çözümü: $\hat{x}_{\mathsf{LS}} = A^\dagger y$

A geniş matris; tam satır kerteli (yani, $(AA^T)^{-1}$ mevcut) ise

$$A^\dagger = A^T(AA^T)^{-1}$$

ile $Ax = b$'nin en küçük norm çözümü: $\hat{x}_{\mathsf{LN}} = A^\dagger y$

SVD'nin kestirme ve evirmede kullanımı

ölçüm modeli: $y = Ax + v$. ölçülen y'den x'i bulmak istiyoruz

- $y \in \mathbb{R}^m$ ölçüm (*measurement*)
- $x \in \mathbb{R}^n$ kestirmek (yani, değerini tahmin etmek) istediğimiz vektör
- $v \in \mathbb{R}^m$ ölçüm gürültüsü veya hata

(not: **kestirme** (*estimation*), **evirme** (*inversion*))

norm-sınırlı (*norm-bound*) gürültü modeli: v için $\|v\| \leq \alpha$ olduğunu varsayıyoruz ancak bunun haricinde v ile ilgili hiçbir şey bilmiyoruz (burada α gürültünün maksimum normu)

SVD'nin kestirme ve evirmede kullanımı

- $\hat{x} = By$ formundaki **kestirici**yi (*estimator*) ele alalım ($BA = I$, yani, kestirici **yansız** (*unbiased*))
- kestirme (veya evirme) hatasına $\tilde{x}$ diyelim:
  $\tilde{x} = \hat{x} - x = \underbrace{By}_{\hat{x}} - \underbrace{B(y - v)}_{x} = Bv$
- olası kestirme hatalarının kümesi

$$\tilde{x} \in \mathcal{E}_{\mathsf{bel}} = \{Bv \mid \|v\| \le \alpha\}$$

  ile verilen bir elipsoittir
- $x = \hat{x} - \tilde{x} \in \hat{x} - \mathcal{E}_{\mathsf{bel}} = \hat{x} + \mathcal{E}_{\mathsf{bel}}$, dolayısıyla: x'in (bilmediğimiz) gerçek değeri, merkezi kestirim $\hat{x}$'te olan **belirsizlik elipsoiti**nin (*uncertainty ellipsoid*) elemanıdır
- iyi kestiricinin belirsizlik elipsoiti $\mathcal{E}_{\mathsf{bel}}$ küçük olur ($BA = I$'yı sağlamak koşuluyla)

SVD'nin kestirme ve evirmede kullanımı

$\mathcal{E}_{\mathsf{bel}}$'nin yarıeksenleri $\alpha\sigma_i u_i$ şeklindedir (yani, B'nin tekil değerleri ve tekil vektörleri)

hatanın normunun maksimum değeri $\alpha\|B\|$, yani $\|\hat{x} - x\| \leq \alpha\|B\|$

en küçük karelerin optimalitesi: $BA = I$ koşulunu sağlayan bütün kestiricileri ele alalım; bunların belirsizlik elipsoitlerine $\mathcal{E}$ diyelim. en küçük kareler kestirici $B_{\mathsf{LS}} = A^\dagger$ olarak verilsin; bunun belirsizlik elipsoitine $\mathcal{E}_{\mathsf{LS}}$ diyelim. buradan

- $B_{\mathsf{LS}} B_{\mathsf{LS}}^T \leq BB^T$
- $\mathcal{E}_{\mathsf{LS}} \subseteq \mathcal{E}$
- $\|B_{\mathsf{LS}}\| \leq \|B\|$

yazabiliriz. sonuç olarak: en küçük kareler kestirici belirsizlik elipsoiti en küçük olan kestiricidir

SVD'nin kontrolde kullanımı

model: $y = Ax$. istenen y'yi oluşturacak x'i seçmek istiyoruz

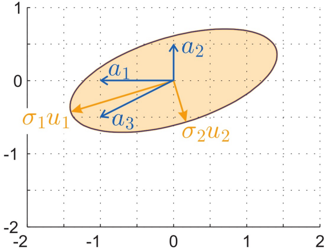
- sağ tekil vektör v_i, tekil değer σ_i ile kuvvetlendirilerek, sol tekil vektör u_i ile eşlenir
- σ_i, u_i yönündeki kontrol otoritesinin bir ölçüsüdür
- $r < m \Rightarrow u_{r+1}, \ldots, u_m$ yönlerinde kontrol otoritesi yok
- A geniş matris; tam satır kerteli (yani, $(AA^T)^{-1}$ mevcut) ise oluşturulabilecek y'lerin kümesi

$$\mathcal{E} = \{y \in \mathbb{R}^m \mid y^T(AA^T)^{-1}y \leq 1\}$$

ile verilen bir elipsoittir

SVD'nin kontrolde kullanımı

örnek: katı cisme uygulanan kuvvetler. örneğin, bir taşıta (araba, uçak, roket vb.) **itki** (*thrust*) sistemleriyle çeşitli yönlerde kuvvetlerin uygulandığı bir uygulamayı ele alalım



$$A = \begin{bmatrix} a_1 & a_2 & a_3 \end{bmatrix}$$
$$= \begin{bmatrix} -1 & 0 & -1 \\ 0 & 0.5 & -0.5 \end{bmatrix}$$
$$\sigma_1 = 1.4668, \quad \sigma_2 = 0.5904$$

- $y = Ax \in \mathbb{R}^m$: cisim üzerindeki toplam kuvvet (birim: N)
- $x_i \in \mathbb{R}$: itki sistemi i'ye sağlanan güç (birim: W)
- $\|a\|_i$: itki sistemi i'nin **etkililiği** (*efficiency*)
- itki uygulayabileceğimiz en etkili yön büyük yarıeksen (yani, u_1) yönüdür

Ana bileşenler analizi

ana bileşenler analizi (*principal component analysis* (PCA)): keşfedici veri çözümlemesi (exploratory data analysis), veri görselleştirme, ve veri ön işleme (data preprocessing) gibi uygulamalarda kullanılan bir doğrusal **boyut indirgeme** (*dimensionality reduction*) yöntemi

örnek: tıbbi veri
örneklem sayısı: 216 (kişi)
öznitelik sayısı: 4000 (gen ifadesi)
veri matrisi: $A \in \mathbb{R}^{216 \times 4000}$
etiketler: {pozitif, negatif}
121 kişi pozitif, 95 kişi negatif
figür: ilk 3 ana bileşen uzayında veri