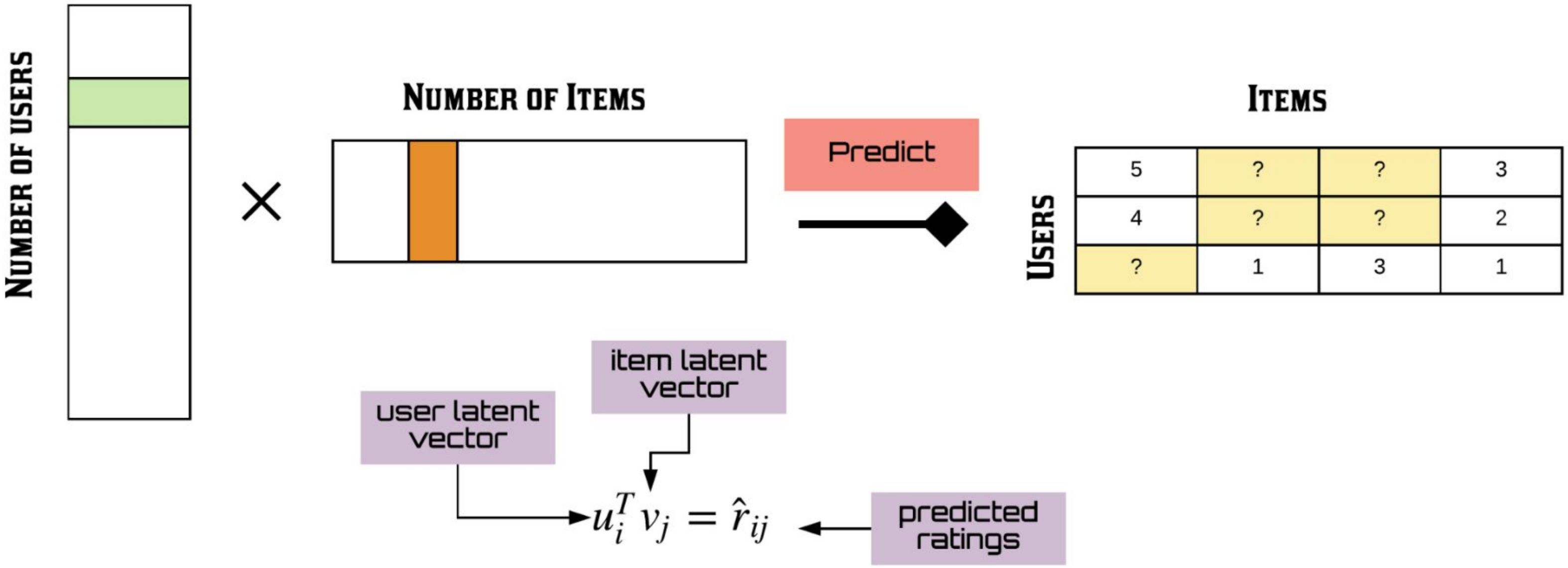
Matris Ayrıştırma

Kamer Kaya
Sabancı Üniversitesi

Yapay Öğrenme için Matematik, 1-6 Şubat 2021, Nesin Matematik Köyü

Sabancı Üniversitesi

Matris Ayrıştırma



Determinant

$$\det(\boldsymbol{A}) = \begin{vmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & \ddots & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{vmatrix}$$

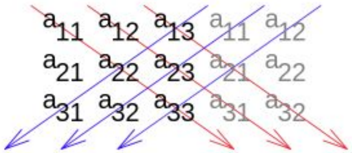
Tanım: Bir **kare** matrisin **determinantı**, o matrisi bir sayıya eşleyen bir fonksiyondur.

$$\det(\boldsymbol{A}) = \det(a_{11}) = a_{11}$$ **(n = 1)**

$$\begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$ **(n = 2)**

$$\begin{vmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{vmatrix} = a_{11}a_{22}a_{33} + a_{21}a_{32}a_{13} + a_{31}a_{12}a_{23}$$

$$- a_{31}a_{22}a_{13} - a_{11}a_{32}a_{23} - a_{21}a_{12}a_{33}$$ **(n = 3)**



Determinant

Teorem: Bir kare matris $\mathbf{A} \in \mathbb{R}^{n \times n}$ 'in tersi ancak ve ancak $det(\mathbf{A}) \neq 0$ ise vardır.

2×2 bir $\boldsymbol{A}$ matrisi için

$$\boldsymbol{A}^{-1} = \frac{1}{a_{11}a_{22} - a_{12}a_{21}} \begin{bmatrix} a_{22} & -a_{12} \\ -a_{21} & a_{11} \end{bmatrix}$$

$$\det(\boldsymbol{A}) = \begin{vmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{vmatrix} = a_{11}a_{22} - a_{12}a_{21}$$

Determinant

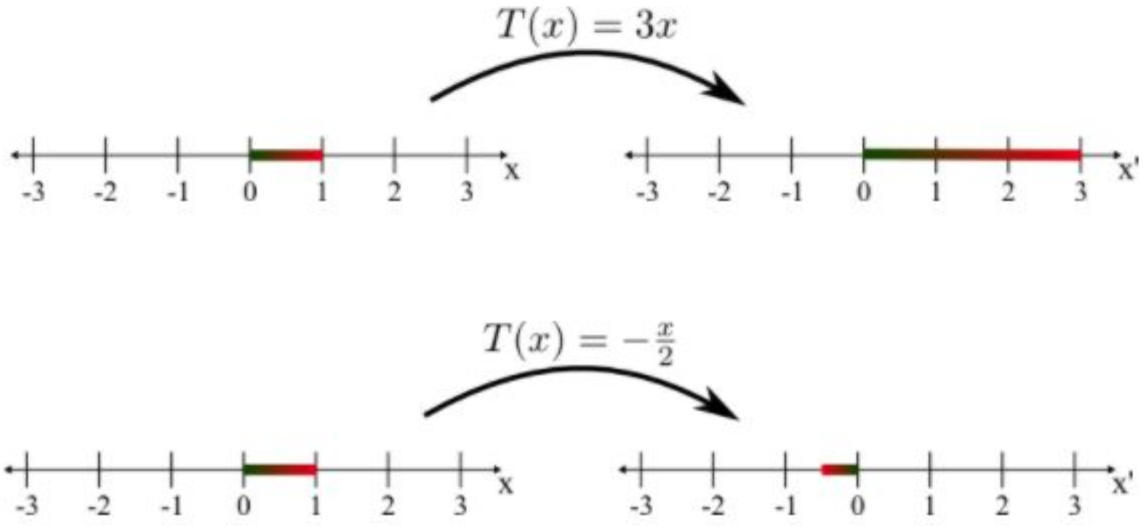
$$L = \begin{bmatrix} \ell_{1,1} & & & & 0 \\ \ell_{2,1} & \ell_{2,2} & & & \\ \ell_{3,1} & \ell_{3,2} & \ddots & & \\ \vdots & \vdots & \ddots & \ddots & \\ \ell_{n,1} & \ell_{n,2} & \dots & \ell_{n,n-1} & \ell_{n,n} \end{bmatrix} \qquad U = \begin{bmatrix} u_{1,1} & u_{1,2} & u_{1,3} & \dots & u_{1,n} \\ & u_{2,2} & u_{2,3} & \dots & u_{2,n} \\ & & \ddots & \ddots & \vdots \\ & & & \ddots & u_{n-1,n} \\ 0 & & & & u_{n,n} \end{bmatrix}$$

Bir **T** matrisinde, **$i > j$** (ya da **$i < j$**) değerleri için eğer **T_{ij} =** 0 ise matrisin determinantı

$$\det(\boldsymbol{T}) = \prod_{i=1}^{n} T_{ii}$$

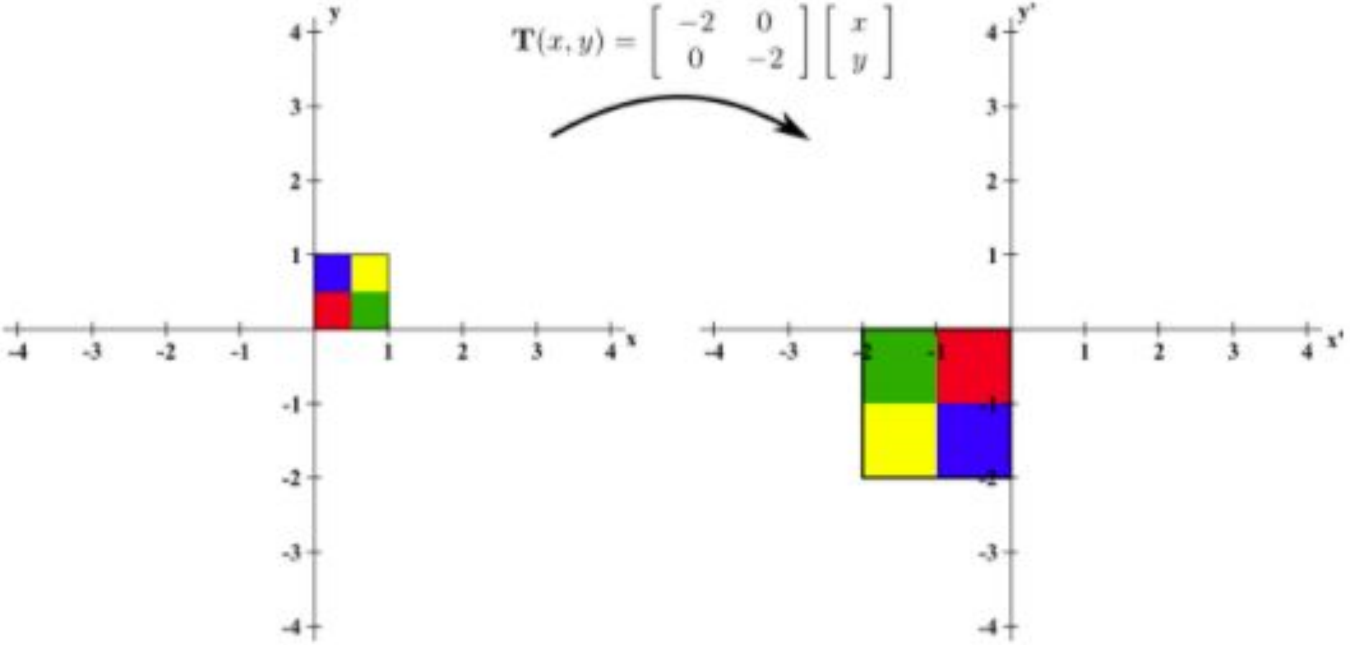
şeklinde hesaplanır.

Determinant; uzunluk, alan, hacim, ...



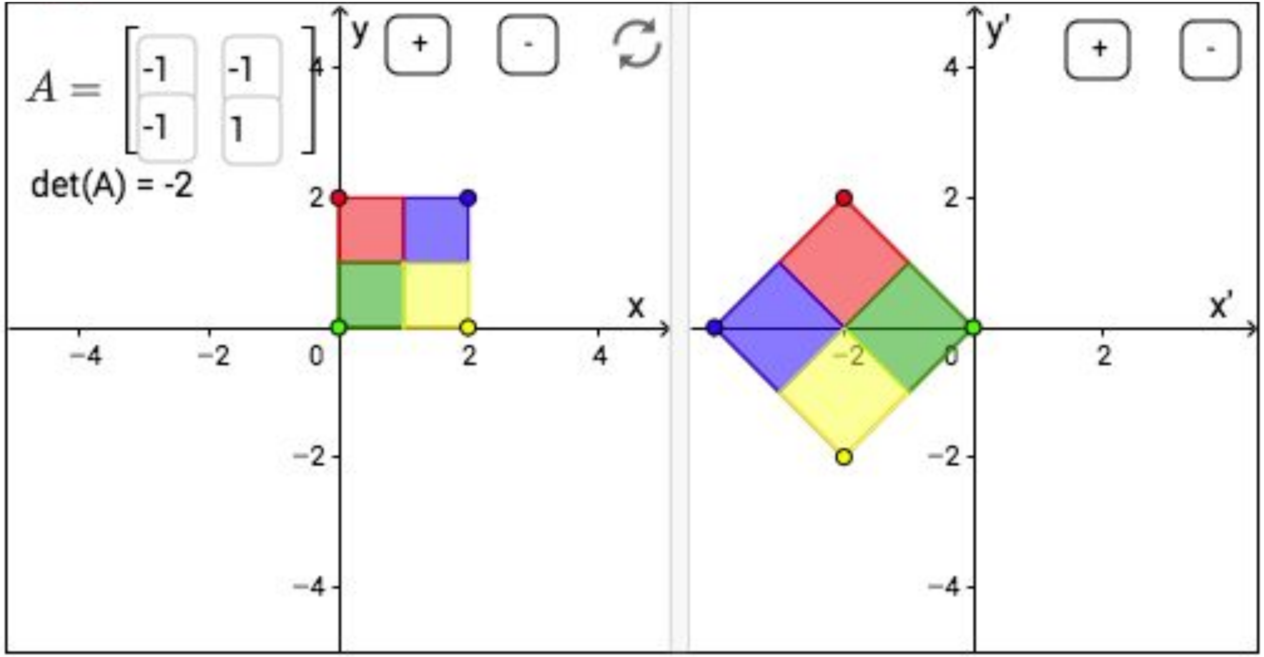
https://mathinsight.org/determinant_linear_transformation

Determinant; **uzunluk**, alan, hacim, ….



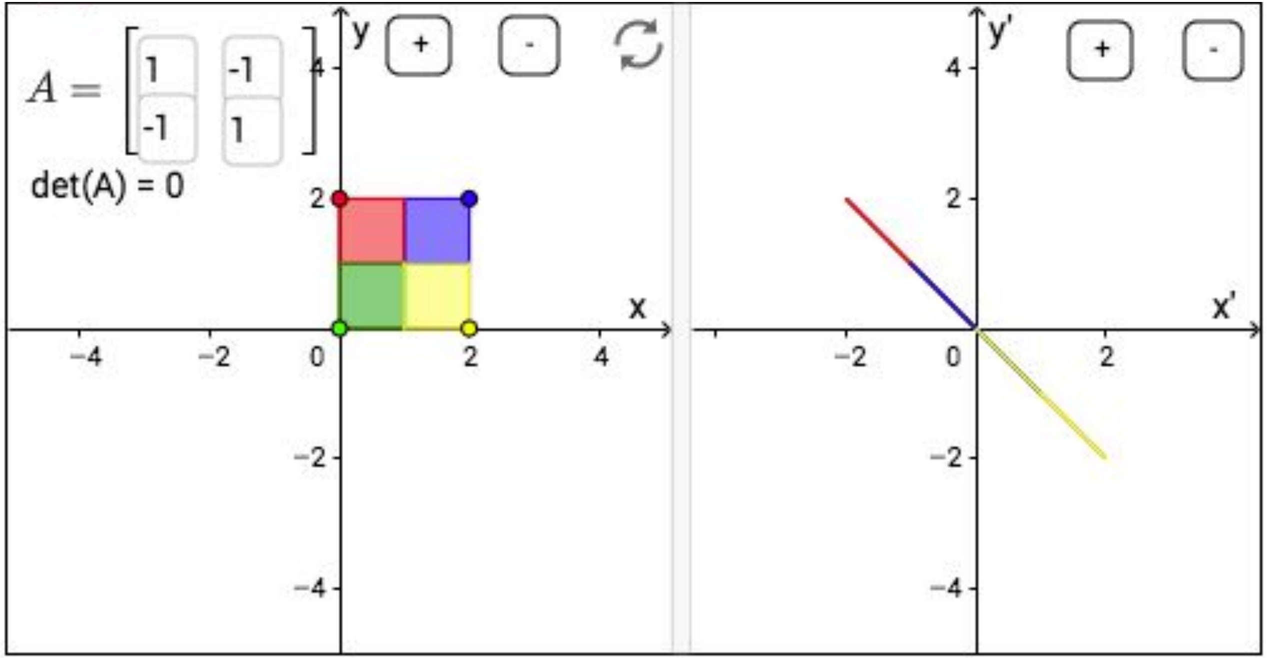
https://mathinsight.org/determinant_linear_transformation

Determinant; **uzunluk**, alan, hacim, ...



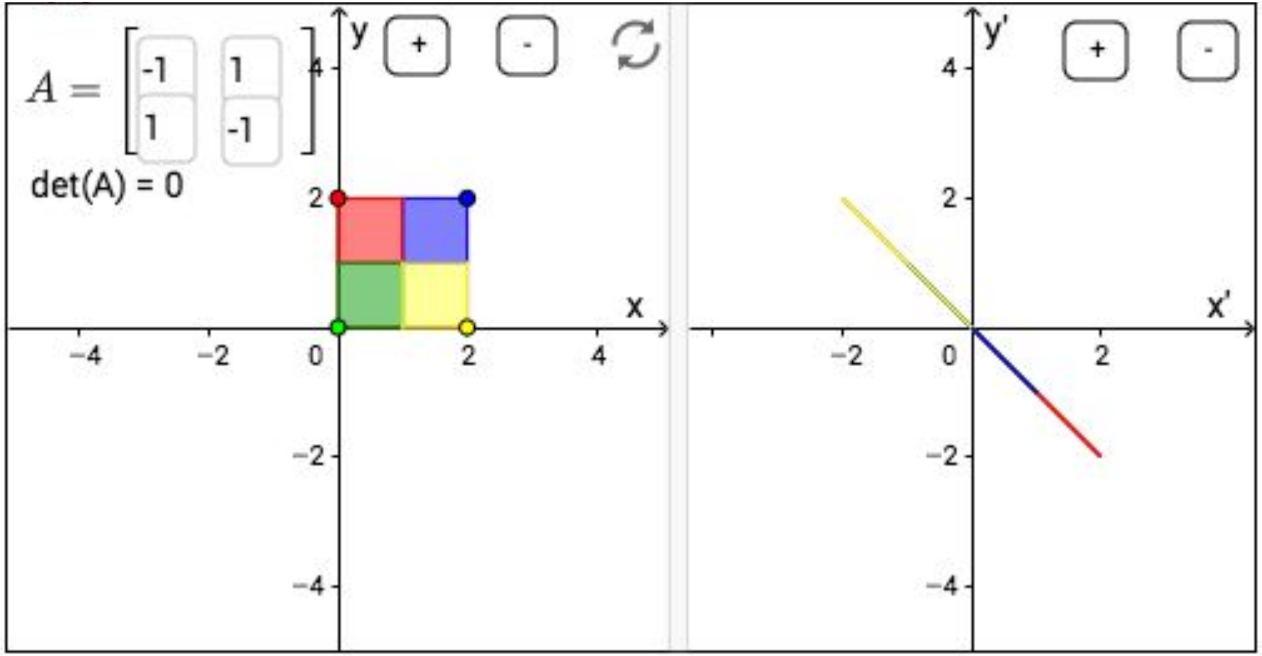
https://mathinsight.org/determinant_linear_transformation

Determinant; **uzunluk**, alan, hacim, ...



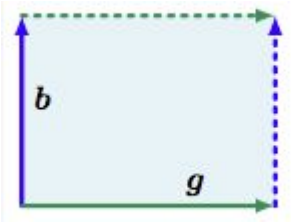
https://mathinsight.org/determinant_linear_transformation

Determinant; **uzunluk**, alan, hacim, ….



https://mathinsight.org/determinant_linear_transformation

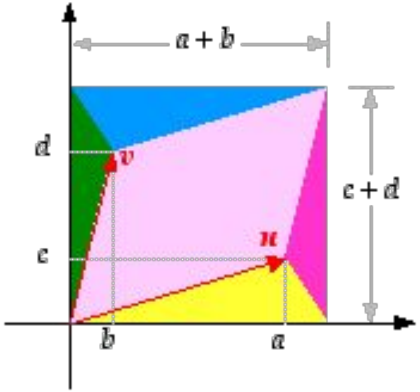
Determinant; uzunluk, alan, hacim, ...



$$\begin{vmatrix} b & 0 \\ 0 & g \end{vmatrix} = bg - 0 = bg.$$

$$\boldsymbol{A} = [\boldsymbol{b}, \boldsymbol{g}]$$

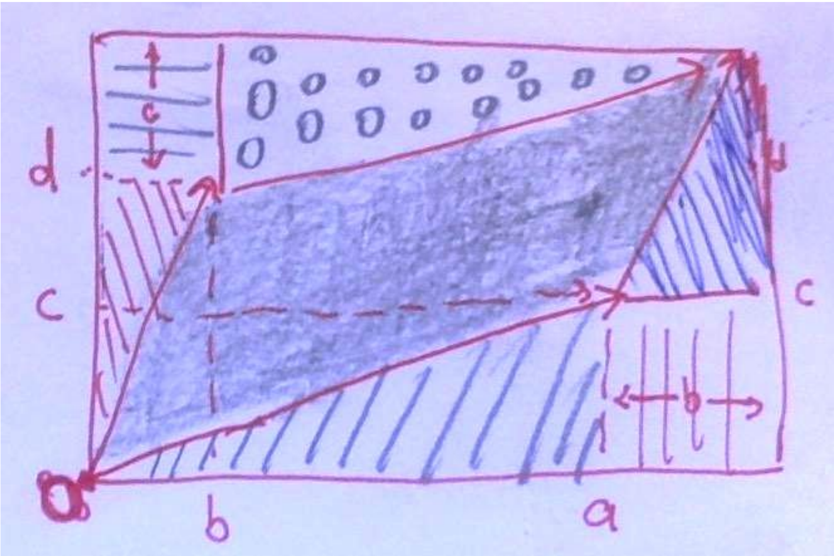
Doğrusal bağımlı **b** ve **g** sütun vektörleri için hem alan hem determinant 0 olur

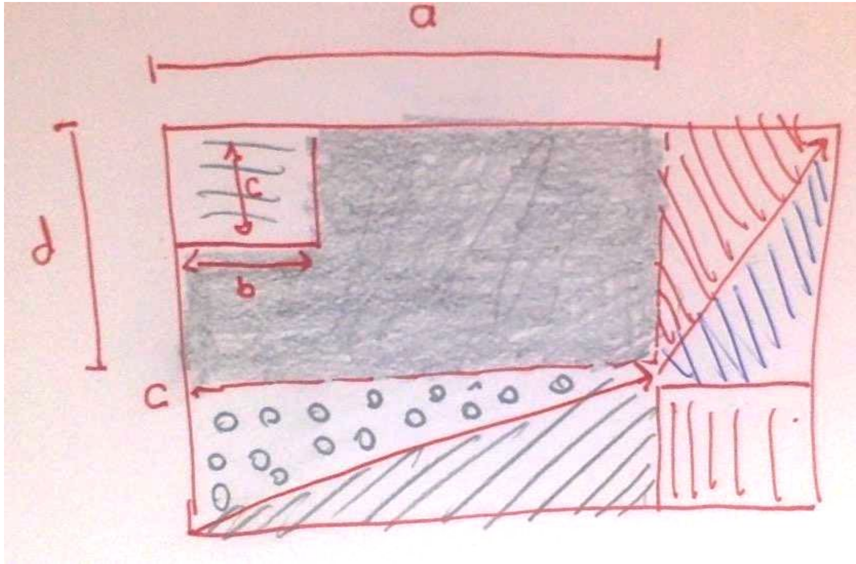


$$\mathbf{A} = \begin{bmatrix} b & a \\ d & c \end{bmatrix}$$

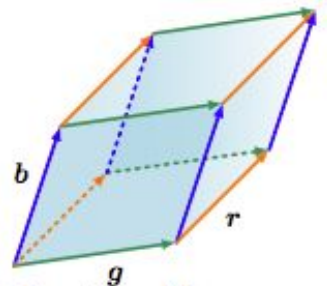
det(**A**) = ((a+b) x (c+d)) - ((a+b) x c + b x (c+d))
ac + ad + bc + bd - ac - bc - bc - bd
-(bc - ad)

Determinant; uzunluk, alan, hacim, ...





Determinant; uzunluk, alan, hacim, ….



$$\boldsymbol{A} = [\boldsymbol{r},\ \boldsymbol{b},\ \boldsymbol{g}]$$

$$\boldsymbol{r} = \begin{bmatrix} 2 \\ 0 \\ -8 \end{bmatrix}, \quad \boldsymbol{g} = \begin{bmatrix} 6 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{b} = \begin{bmatrix} 1 \\ 4 \\ -1 \end{bmatrix}$$

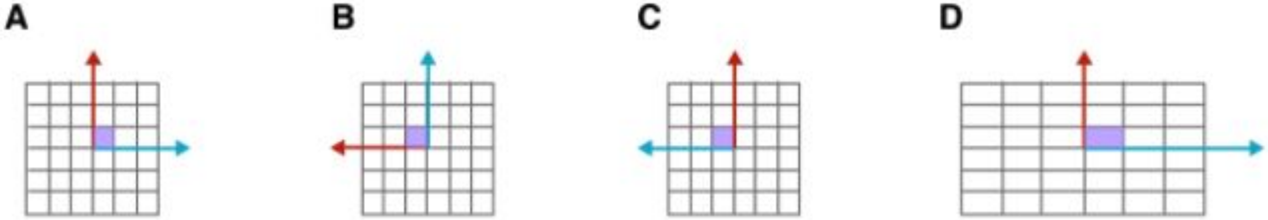
$$\boldsymbol{A} = [\boldsymbol{r},\ \boldsymbol{g},\ \boldsymbol{b}] = \begin{bmatrix} 2 & 6 & 1 \\ 0 & 1 & 4 \\ -8 & 0 & -1 \end{bmatrix}$$

$$V = |\det(\boldsymbol{A})| = 186.$$

Determinant

$$\begin{bmatrix} \alpha_1 & & & & \\ & \ddots & & & \\ & & \alpha_i & & \\ & & & \ddots & \\ & & & & \alpha_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} \alpha_1 \\ \\ \\ \\ \\ \end{bmatrix} + \cdots + x_i \begin{bmatrix} \\ \\ \alpha_i \\ \\ \\ \end{bmatrix} + \cdots + x_n \begin{bmatrix} \\ \\ \\ \\ \alpha_n \end{bmatrix}$$

Köşegensel matrisler

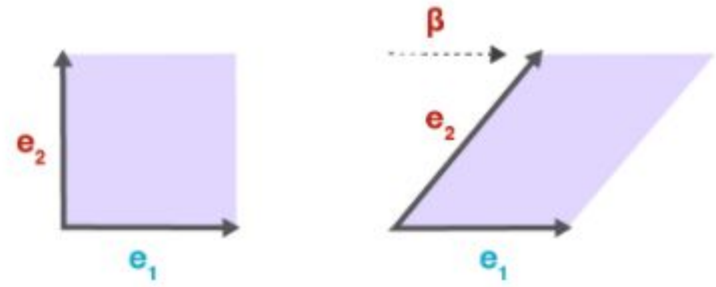


http://gregorygundersen.com/blog/2018/10/24/matrices/

Determinant

$$\begin{bmatrix} \alpha_1 & \cdots & \beta & & \\ & \ddots & \vdots & & \\ & & \alpha_i & & \\ & & & \ddots & \\ & & & & \alpha_n \end{bmatrix} \begin{bmatrix} x_1 \\ \vdots \\ x_i \\ \vdots \\ x_n \end{bmatrix} = x_1 \begin{bmatrix} \alpha_1 \\ \\ \\ \\ \\ \end{bmatrix} + \cdots + x_i \begin{bmatrix} \beta \\ \\ \alpha_i \\ \\ \\ \end{bmatrix} + \cdots + x_n \begin{bmatrix} \\ \\ \\ \\ \alpha_n \end{bmatrix}$$

Kayma matrisleri



http://gregorygundersen.com/blog/2018/10/24/matrices/

Determinant

Ortogonal matris, sütun vektörleri birim vektörler ve birbirlerine dik olan kare matrislerdir. İki vektör birbirine dik ise iç çarpımları 0'dır.

Ortogonal matrisler

Ortogonal olmayan vektörler

Ortogonal olan vektörler

$$A^{\top}A = \begin{bmatrix} \mathbf{a}_1^{\top} \\ \hline \mathbf{a}_2^{\top} \\ \hline \vdots \\ \hline \mathbf{a}_n^{\top} \end{bmatrix} \left[\ \mathbf{a}_1 \mid \mathbf{a}_2 \mid \dots \mid \mathbf{a}_n \ \right] = \begin{bmatrix} 1 & & & \\ & 1 & & \\ & & \ddots & \\ & & & 1 \end{bmatrix} = I_n$$

Determinant

- Ortogonal matrislerin sağladığı dönüşüm vektörler arasındaki açıyı korur

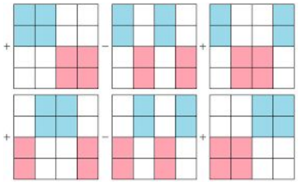
$$\begin{aligned} A\mathbf{v}_i \cdot A\mathbf{v}_j &= (A\mathbf{v}_i)^\top (A\mathbf{v}_j) \\ &= \mathbf{v}_i^\top A^\top A \mathbf{v}_j \\ &= \mathbf{v}_i^\top \mathbf{v}_j \\ &= \mathbf{v}_i \cdot \mathbf{v}_j \end{aligned}$$

- Ortogonal matrislerin sağladığı dönüşüm vektörlerin uzunluklarını korur

$$\begin{aligned} \|A\mathbf{v}\|_2^2 &= A\mathbf{v} \cdot A\mathbf{v} \\ &= (A\mathbf{v})^\top (A\mathbf{v}) \\ &= \mathbf{v}^\top A^\top A \mathbf{v} \\ &= \mathbf{v}^\top \mathbf{v} \\ &= \mathbf{v} \cdot \mathbf{v} \\ &= \|\mathbf{v}\|_2^2 \end{aligned}$$

- Ortogonal matrislerin determinantları 1 ya da -1'dir

$$\begin{aligned} 1 &= \det(I_n) = \det(A^\top A) \\ &= \det(A^\top)\det(A) = (\det(A))^2 \end{aligned}$$



Determinant

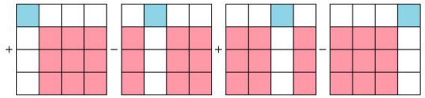
Teorem: Bir $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ matrisimiz olsun. Bütün $j = 1, \ldots, n$ değerleri için:

1) **j** sütunundan yapılan **Laplace açılımı** ile

$$\det(\boldsymbol{A}) = \sum_{k=1}^{n} (-1)^{k+j} a_{kj} \det(\boldsymbol{A}_{k,j})$$

2) **j** satırından yapılan Laplace açılımı ile

$$\det(\boldsymbol{A}) = \sum_{k=1}^{n} (-1)^{k+j} a_{jk} \det(\boldsymbol{A}_{j,k})$$



Burada $\boldsymbol{A}_{k,j} \in \mathbb{R}^{(n-1)\times(n-1)}$ satır **k** ve sütun **j**'yi silerek elde edilen alt-matrisi ifade eder.

https://www.geometrictools.com/Documentation/LaplaceExpansionTheorem.pdf

Determinant

Örnek: $\boldsymbol{A} = \begin{bmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 0 & 1 \end{bmatrix}$ matrisi için

ilk satır üzerinden yapılan Laplace açılımı ile

$$\begin{vmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 0 & 0 & 1 \end{vmatrix} = (-1)^{1+1} \cdot 1 \begin{vmatrix} 1 & 2 \\ 0 & 1 \end{vmatrix} + (-1)^{1+2} \cdot 2 \begin{vmatrix} 3 & 2 \\ 0 & 1 \end{vmatrix} + (-1)^{1+3} \cdot 3 \begin{vmatrix} 3 & 1 \\ 0 & 0 \end{vmatrix}$$

$$\det(\boldsymbol{A}) = 1(1-0) - 2(3-0) + 3(0-0) = -5$$

Ya da $\det(\boldsymbol{A}) = 1{\cdot}1{\cdot}1 + 3{\cdot}0{\cdot}3 + 0{\cdot}2{\cdot}2 - 0{\cdot}1{\cdot}3 - 1{\cdot}0{\cdot}2 - 3{\cdot}2{\cdot}1 = 1 - 6 = -5$

Determinant

- Matrisin iki satırı ya da iki sütunu yer değiştirirse determinant -1 ile çarpılır.
- Matrisin bir satırı (sütunu) başka bir satıra (sütuna) eklenirse determinant değişmez.
- Bir **A** matrisinin tersi varsa $\det(\boldsymbol{A}^{-1}) = \frac{1}{\det(\boldsymbol{A})}$

$$\det(\lambda \boldsymbol{A}) = \lambda^n \det(\boldsymbol{A})$$

$$\det(\boldsymbol{AB}) = \det(\boldsymbol{A})\det(\boldsymbol{B})$$

$$\det(\boldsymbol{A}) = \det(\boldsymbol{A}^\top)$$

- $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ kare matrisinin determinantı ancak ve ancak mertebesi **n** ise 0'dan farklıdır.

İz (Trace)

Bir $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ matrisinin **izi**, köşegenindeki değerlerin toplamına eşittir:

$$\operatorname{tr}(\boldsymbol{A}) := \sum_{i=1}^{n} a_{ii}$$

İz'in şu özellikleri vardır

- $\operatorname{tr}(\boldsymbol{A}+\boldsymbol{B}) = \operatorname{tr}(\boldsymbol{A}) + \operatorname{tr}(\boldsymbol{B}) \qquad \boldsymbol{A}, \boldsymbol{B} \in \mathbb{R}^{n \times n}$
- $\operatorname{tr}(\alpha\boldsymbol{A}) = \alpha\operatorname{tr}(\boldsymbol{A}),\, \alpha \in \mathbb{R} \qquad \boldsymbol{A} \in \mathbb{R}^{n \times n}$
- $\operatorname{tr}(\boldsymbol{I}_n) = n$
- $\operatorname{tr}(\boldsymbol{AB}) = \operatorname{tr}(\boldsymbol{BA}) \qquad \boldsymbol{A} \in \mathbb{R}^{n \times k}, \boldsymbol{B} \in \mathbb{R}^{k \times n}$

İz

İz fonksiyonu döngüsel permütasyonlar altında değişmezdir

$$\mathrm{tr}(\boldsymbol{AKL}) = \mathrm{tr}(\boldsymbol{KLA}) \qquad \boldsymbol{A} \in \mathbb{R}^{a\times k}, \boldsymbol{K} \in \mathbb{R}^{k\times l}, \boldsymbol{L} \in \mathbb{R}^{l\times a}$$

$$\mathrm{tr}(\boldsymbol{x}\boldsymbol{y}^\top) = \mathrm{tr}(\boldsymbol{y}^\top\boldsymbol{x}) = \boldsymbol{y}^\top\boldsymbol{x} \in \mathbb{R} \qquad \boldsymbol{x}, \boldsymbol{y} \in \mathbb{R}^n$$

Bir ***V*** vektör uzayı ve $\Phi : V \to V$, eşlemesi için dönüşüm matrisi **A** olsun. Farklı bir baz kullanıldığında ise dönüşüm matrisi **B** = $\mathbf{S}^{-1}\mathbf{AS}$ olsun.

$$\mathrm{tr}(\boldsymbol{B}) = \mathrm{tr}(\boldsymbol{S}^{-1}\boldsymbol{AS}) = \mathrm{tr}(\boldsymbol{ASS}^{-1}) = \mathrm{tr}(\boldsymbol{A})$$

olur. Doğrusal dönüşümlerin matrisleri baz ile değişse de dönüşümün izi farklı baz kullanıldığında değişmez.

Karakteristik Polinom

Bir $\lambda \in \mathbb{R}$ ve bir kare matris $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ için **karakteristik polinom**

$$\begin{aligned} p_{\boldsymbol{A}}(\lambda) &:= \det(\boldsymbol{A} - \lambda \boldsymbol{I}) \\ &= c_0 + c_1\lambda + c_2\lambda^2 + \cdots + c_{n-1}\lambda^{n-1} + (-1)^n\lambda^n \end{aligned}$$

$c_0, \ldots, c_{n-1} \in \mathbb{R}$ şeklinde gösterilir. Burada şu eşitlikler önemlidir.

$$\begin{aligned} c_0 &= \det(\boldsymbol{A}), \\ c_{n-1} &= (-1)^{n-1}\operatorname{tr}(\boldsymbol{A}) \end{aligned}$$

Özdeğer (eigenvalue) ve özvektör (eigenvector)

$\boldsymbol{A} \in \mathbb{R}^{n \times n}$ kare bir matris olsun:

$$\boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x}$$

eşitliğini sağlayan $\lambda \in \mathbb{R}$ değerine A matrisinin bir ***özdeğeri***, buna eşlenik $\boldsymbol{x} \in \mathbb{R}^n \backslash \{\boldsymbol{0}\}$ vektörüne de $\boldsymbol{A}$ matrisinin bir ***özvektörü*** denir. Aşağıdaki ifadeler eşdeğerdir:

- $\lambda \in \mathbb{R}$ değeri $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ matrisinin bir özdeğeridir.
- $(\boldsymbol{A} - \lambda \boldsymbol{I}_n)\boldsymbol{x} = \boldsymbol{0}$ ifadesinin **0**'dan farklı bir $\boldsymbol{x} \neq \boldsymbol{0}$ çözümü vardır.
- $\mathrm{rk}(\boldsymbol{A} - \lambda \boldsymbol{I}_n) < n$
- $\det(\boldsymbol{A} - \lambda \boldsymbol{I}_n) = 0$

Özdeğerler ve özvektörler

İki vektör aynı yönü gösteriyorsa bu vektörler **yöndeştir**. Eğer vektörler aynı ya da zıt yönleri gösteriyorlarsa bu vektörler **doğrusaldır**.

Eğer $\boldsymbol{x}$ vektörü $\boldsymbol{A}$ matrisinin λ özdeğerine tekabül eden bir özvektörü ise, bir $c \in \mathbb{R} \backslash \{0\}$ pozitif reel sayısı için $c\boldsymbol{x}$ vektörü de $\boldsymbol{A}$ matrisinin bir özvektörüdür.

$$\boldsymbol{A}(c\boldsymbol{x}) = c\boldsymbol{A}\boldsymbol{x} = c\lambda\boldsymbol{x} = \lambda(c\boldsymbol{x})$$

Dolayısıyla, $\boldsymbol{x}$ ile doğrusal bütün vektörler $\boldsymbol{A}$ matrisinin bir özvektörüdür.

Özdeğerler ve özvektörler

Teorem: Bir $\lambda \in \mathbb{R}$ reel sayısı ancak ve ancak $p_{\boldsymbol{A}}(\lambda)$ karakteristik polinomunun bir kökü ise $\boldsymbol{A}$ matrisinin bir özdeğeridir.

λ değerinin $p_{\boldsymbol{A}}(\lambda)$ polinomunun kökleri arasında görülme sayısına λ özdeğerinin ***(cebirsel) katlılığı*** denir.

Bir $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ matrisi için λ özdeğerine tekabül eden bütün özvektörlerin oluşturduğu uzay $\mathbb{R}^n$ uzayının bir alt uzayıdır. Bu alt uzaya $\boldsymbol{A}$ matrisinin λ özdeğerine göre **özuzayı** denir ve ile E_λ gösterilir.

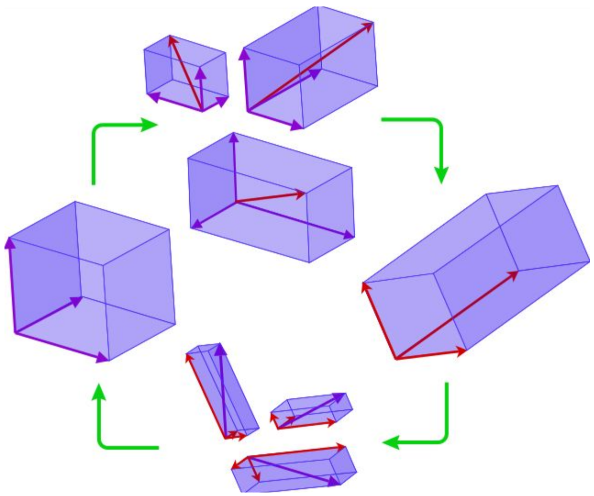
$\boldsymbol{A}$ matrisinin bütün özdeğerlerinin oluşturduğu kümeye, matrisin **özspektrumu** denir.

Eğer $\boldsymbol{A}$ birim matris ise tek özdeğeri (tekrar sayısı n olan) 1'dir. Bütün $\boldsymbol{x} \in \mathbb{R}^n \backslash \{\boldsymbol{0}\}$ vektörleri bir özvektördür. Dolayısıyla E_1 uzayı, $\mathbb{R}^n$ uzayının standart baz vektörlerinin kapsadığı alt uzaydır.

Özdeğerler ve özvektörler

Tanım: $\boldsymbol{A}, \tilde{\boldsymbol{A}} \in \mathbb{R}^{m\times n}$ matrisleri $\tilde{\boldsymbol{A}} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{S}$ eşitliğini sağlayan düzenli (tersi alınabilen) $\boldsymbol{S} \in \mathbb{R}^{n\times n}$ ve $\boldsymbol{T} \in \mathbb{R}^{m\times m}$ matrisleri bulunabiliyorsa birbirine **denktir**.

Tanım: $\boldsymbol{A}, \tilde{\boldsymbol{A}} \in \mathbb{R}^{n\times n}$ matrisleri $\tilde{\boldsymbol{A}} = \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$ eşitliğini sağlayan düzenli (tersi alınabilen) bir $\boldsymbol{S} \in \mathbb{R}^{n\times n}$ matrisi bulunabiliyorsa birbirine **benzerdir**.



https://en.wikipedia.org/wiki/Change_of_basis

Özdeğerler ve özvektörler

- Bir matrisin kendisi ve devriği aynı özdeğerlere sahiptir.
    - Fakat özvektörleri farklı olabilir.
- Bir E_λ alt uzayı $\boldsymbol{A} - \lambda \boldsymbol{I}$ 'in sıfır uzayına eşittir.

$$\begin{aligned} \boldsymbol{A}\boldsymbol{x} = \lambda \boldsymbol{x} \iff & \boldsymbol{A}\boldsymbol{x} - \lambda \boldsymbol{x} = \boldsymbol{0} \\ \iff & (\boldsymbol{A} - \lambda \boldsymbol{I})\boldsymbol{x} = \boldsymbol{0} \iff \boldsymbol{x} \in \ker(\boldsymbol{A} - \lambda \boldsymbol{I}) \end{aligned}$$

- Benzer matrislerin **özspektrumu** aynıdır. Bu da doğrusal bir eşleme Φ'nin özdeğerlerinin dönüşüm matrisi için kullanılan baz vektörlerinden bagımsız olduğunu gösterir. Dolayısıyla özdeğerler, determinant ve iz bir eşleme için baz değişikliği durumlarında değişmez
- Simetrik ve **pozitif kesin** (kesin artı) matrisler her zaman pozitif ve reel özdeğerlere sahiptir.

$$\mathbf{x}^T \mathbf{A} \mathbf{x} > 0 \text{ for } \mathbf{x} \neq \mathbf{0}$$

Özdeğerler ve özvektörler

Örnek: $\boldsymbol{A} = \begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix}$ matrisinin özdeğerleri şu şekilde bulunur.

$$\begin{aligned} p_{\boldsymbol{A}}(\lambda) &= \det(\boldsymbol{A} - \lambda \boldsymbol{I}) \\ &= \det\left(\begin{bmatrix} 4 & 2 \\ 1 & 3 \end{bmatrix} - \begin{bmatrix} \lambda & 0 \\ 0 & \lambda \end{bmatrix}\right) = \begin{vmatrix} 4-\lambda & 2 \\ 1 & 3-\lambda \end{vmatrix} \\ &= (4-\lambda)(3-\lambda) - 2 \cdot 1 . \end{aligned}$$

$$p(\lambda) = (4-\lambda)(3-\lambda) - 2 \cdot 1 = 10 - 7\lambda + \lambda^2 = (2-\lambda)(5-\lambda)$$

$$\lambda_1 = 2 \quad \lambda_2 = 5.$$

Özdeğerler ve özvektörler

Özvektörler ve özuzay hesabı şu şekilde yapılabilir

$$\begin{bmatrix} 4-\lambda & 2 \\ 1 & 3-\lambda \end{bmatrix} \boldsymbol{x} = \mathbf{0}$$

$$\mathbf{A}x = \lambda\mathbf{x}$$
$$\mathbf{A}x = \mathbf{I}\lambda\mathbf{x}$$
$$(\mathbf{A} - \mathbf{I}\lambda)\mathbf{x} = 0$$

$(\lambda = 5)$

$$\begin{bmatrix} 4-5 & 2 \\ 1 & 3-5 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -1 & 2 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0}$$

$$E_5 = \text{span}[\begin{bmatrix} 2 \\ 1 \end{bmatrix}]$$

Özdeğerler ve özvektörler

Özvektörler ve özuzay hesabı şu şekilde yapılabilir

$$\mathbf{A}x = \lambda\mathbf{x}$$
$$\mathbf{A}x = \mathbf{I}\lambda\mathbf{x}$$
$$(\mathbf{A} - \mathbf{I}\lambda)\mathbf{x} = 0$$

$$\begin{bmatrix} 4-\lambda & 2 \\ 1 & 3-\lambda \end{bmatrix} \boldsymbol{x} = \mathbf{0}$$

$$(\lambda = 2) \qquad \begin{bmatrix} 4-2 & 2 \\ 1 & 3-2 \end{bmatrix} \boldsymbol{x} = \begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix} \boldsymbol{x} = \mathbf{0}$$

$$E_2 = \text{span}[\begin{bmatrix} 1 \\ -1 \end{bmatrix}]$$

Özdeğerler ve özvektörler

Örnek: $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 4 & 9 \end{bmatrix}$

$$|A - \lambda I| = \left| \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 4 & 9 \end{bmatrix} - \lambda \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \right| = \begin{vmatrix} 2-\lambda & 0 & 0 \\ 0 & 3-\lambda & 4 \\ 0 & 4 & 9-\lambda \end{vmatrix}$$

$$= (2-\lambda)\left[(3-\lambda)(9-\lambda) - 16\right] = -\lambda^3 + 14\lambda^2 - 35\lambda + 22$$

Özdeğerler	2	1	11
	$[1 \quad 0 \quad 0]^T$	$[0 \quad -2 \quad 1]^T$	$[0 \quad 1 \quad 2]^T$

Özdeğerler ve özvektörler

Örnek: $A = \begin{bmatrix} 2 & 0 & 0 \\ 0 & 3 & 4 \\ 0 & 4 & 9 \end{bmatrix}$

$$\begin{bmatrix} 2-\lambda & 0 & 0 \\ 0 & 3-\lambda & 4 \\ 0 & 4 & 9-\lambda \end{bmatrix} \mathbf{x} = 0$$

Özdeğerler:	2	1	11
	$\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 4 \\ 0 & 4 & 7 \end{bmatrix} \mathbf{x} = \mathbf{0}$	$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 4 \\ 0 & 4 & 8 \end{bmatrix} \mathbf{x} = \mathbf{0}$	$\begin{bmatrix} -9 & 0 & 0 \\ 0 & -8 & 4 \\ 0 & 4 & -2 \end{bmatrix} \mathbf{x} = \mathbf{0}$
Özvektörler:	$\begin{bmatrix} 1 & 0 & 0 \end{bmatrix}^T$	$\begin{bmatrix} 0 & -2 & 1 \end{bmatrix}^T$	$\begin{bmatrix} 0 & 1 & 2 \end{bmatrix}^T$

Özdeğerler ve özvektörler

λ_i değeri kare $\boldsymbol{A}$ matrisinin bir özdeğeri olsun. λ_i 'nin **geometrik katlılığı**, ona karşılık gelen, doğrusal bağımsız özvektörlerin sayısına eşittir.

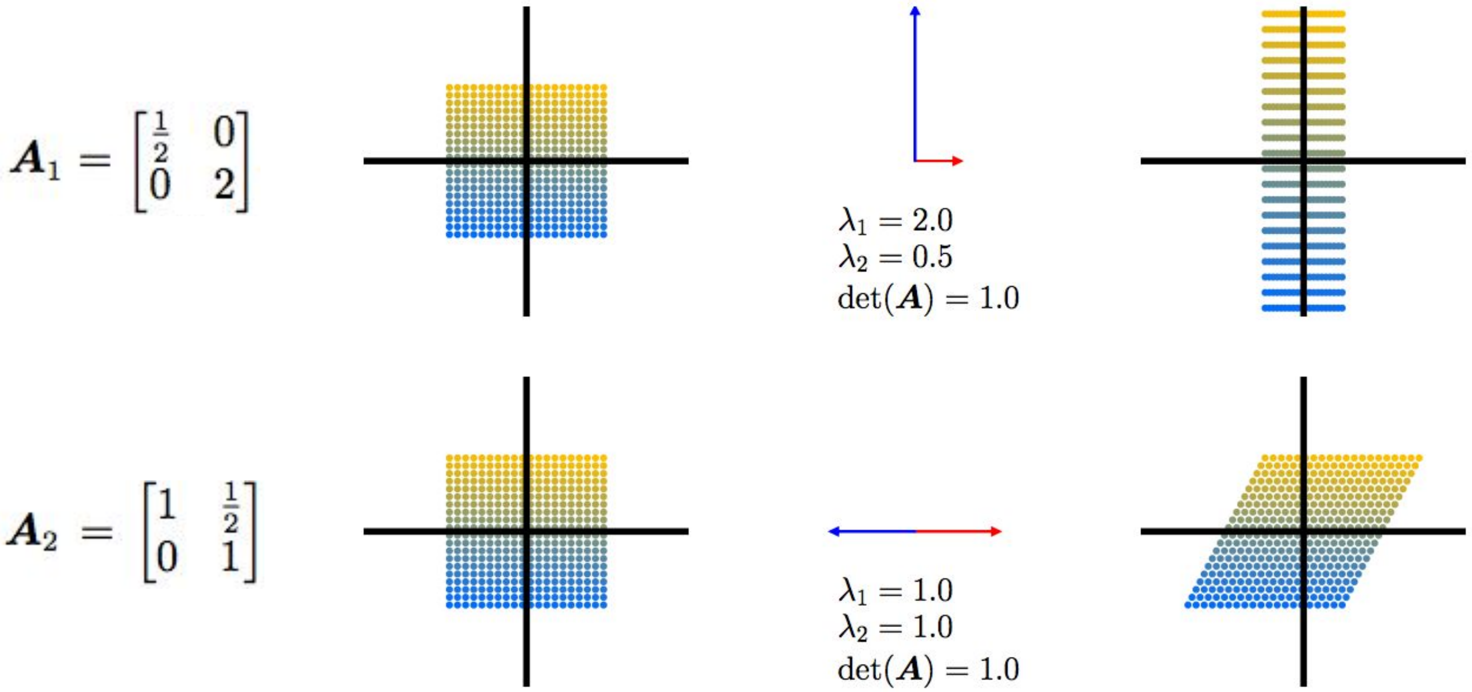
Bir özdeğerin geometrik katlılığı onun aritmetik katlılığınından küçük eşittir.

Örnek:

$$\boldsymbol{A} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \qquad \lambda_1 = \lambda_2 = 2 \qquad \boldsymbol{x}_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

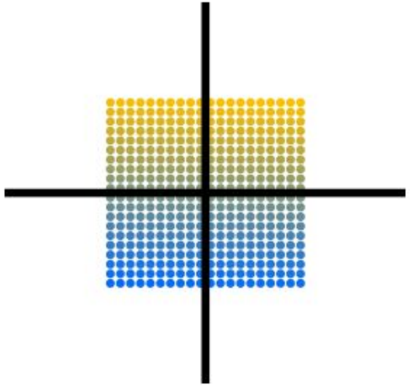
$$\begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \mathbf{x} = \mathbf{0}$$

Özdeğerler ve özvektörler

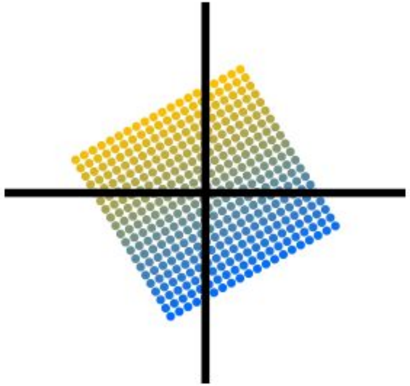


Özdeğerler ve özvektörler

$$\boldsymbol{A}_3 = \begin{bmatrix} \cos(\frac{\pi}{6}) & -\sin(\frac{\pi}{6}) \\ \sin(\frac{\pi}{6}) & \cos(\frac{\pi}{6}) \end{bmatrix} = \frac{1}{2}\begin{bmatrix} \sqrt{3} & -1 \\ 1 & \sqrt{3} \end{bmatrix}$$

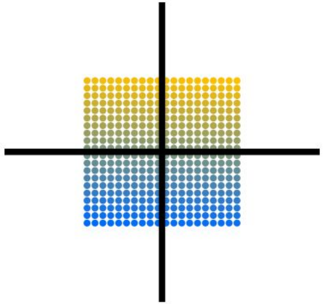


$\lambda_1 = (0.87\text{-}0.5j)$
$\lambda_2 = (0.87{+}0.5j)$
$\det(\boldsymbol{A}) = 1.0$

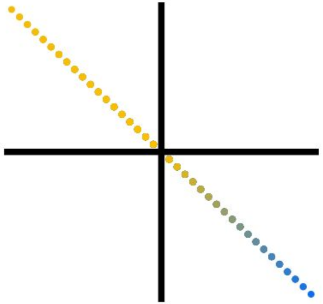


Özdeğerler ve özvektörler

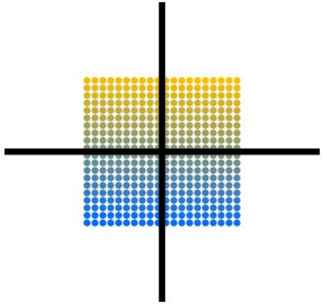
$$\boldsymbol{A}_4 = \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix}$$



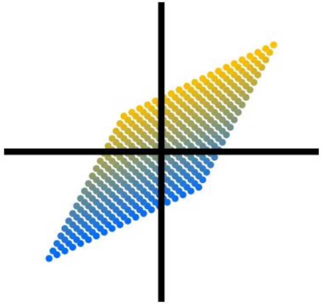
$\lambda_1 = 0.0$
$\lambda_2 = 2.0$
$\det(\boldsymbol{A}) = 0.0$



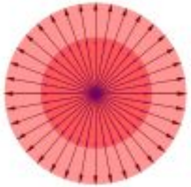
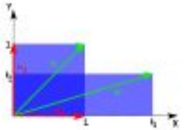
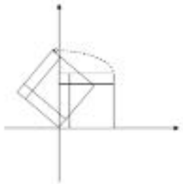
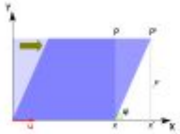
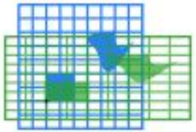
$$\boldsymbol{A}_5 = \begin{bmatrix} 1 & \frac{1}{2} \\ \frac{1}{2} & 1 \end{bmatrix}$$



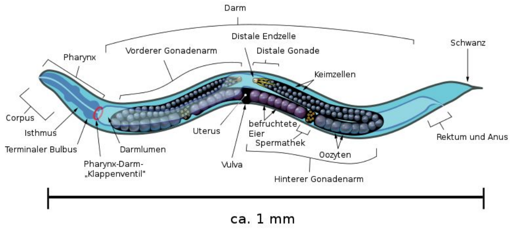
$\lambda_1 = 0.5$
$\lambda_2 = 1.5$
$\det(\boldsymbol{A}) = 0.75$



https://en.wikipedia.org/wiki/Eigenvalues_and_eigenvectors

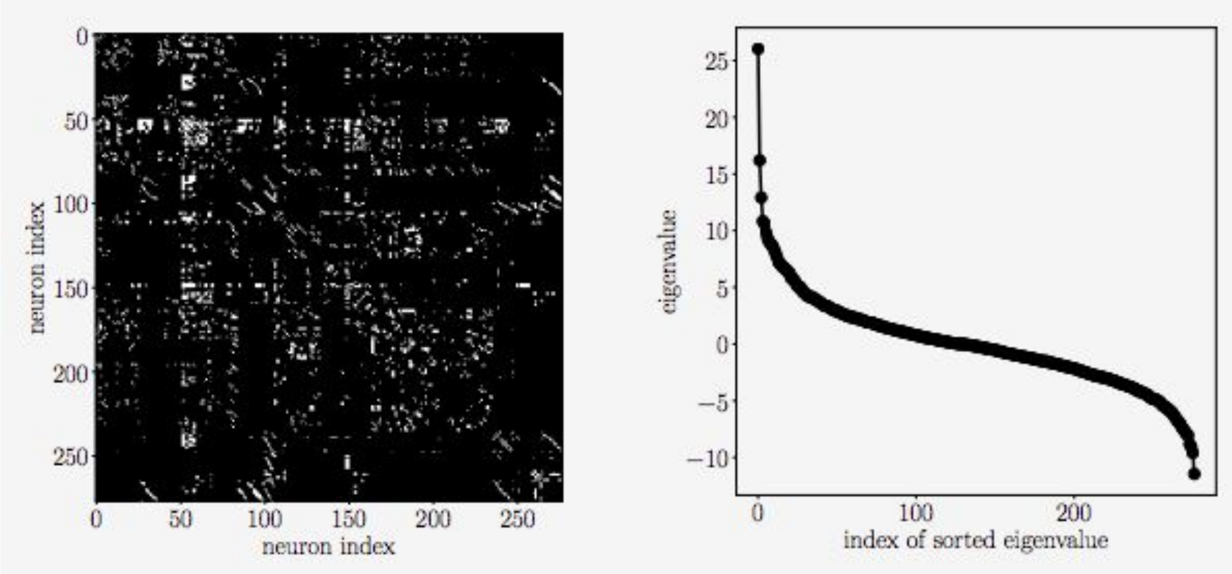
	Scaling	Unequal scaling	Rotation	Horizontal shear	Hyperbolic rotation
Illustration					
Matrix	$\begin{bmatrix} k & 0 \\ 0 & k \end{bmatrix}$	$\begin{bmatrix} k_1 & 0 \\ 0 & k_2 \end{bmatrix}$	$\begin{bmatrix} c & -s \\ s & c \end{bmatrix}$ $c = \cos\theta$ $s = \sin\theta$	$\begin{bmatrix} 1 & k \\ 0 & 1 \end{bmatrix}$	$\begin{bmatrix} c & s \\ s & c \end{bmatrix}$ $c = \cosh\varphi$ $s = \sinh\varphi$
Characteristic polynomial	$(\lambda - k)^2$	$(\lambda - k_1)(\lambda - k_2)$	$\lambda^2 - 2c\lambda + 1$	$(\lambda - 1)^2$	$\lambda^2 - 2c\lambda + 1$
Eigenvalues, λ_i	$\lambda_1 = \lambda_2 = k$	$\lambda_1 = k_1$ $\lambda_2 = k_2$	$\lambda_1 = e^{i\theta} = c + s\mathbf{i}$ $\lambda_2 = e^{-i\theta} = c - s\mathbf{i}$	$\lambda_1 = \lambda_2 = 1$	$\lambda_1 = e^{\varphi}$ $\lambda_2 = e^{-\varphi}$,
Algebraic mult., $\mu_i = \mu(\lambda_i)$	$\mu_1 = 2$	$\mu_1 = 1$ $\mu_2 = 1$	$\mu_1 = 1$ $\mu_2 = 1$	$\mu_1 = 2$	$\mu_1 = 1$ $\mu_2 = 1$
Geometric mult., $\gamma_i = \gamma(\lambda_i)$	$\gamma_1 = 2$	$\gamma_1 = 1$ $\gamma_2 = 1$	$\gamma_1 = 1$ $\gamma_2 = 1$	$\gamma_1 = 1$	$\gamma_1 = 1$ $\gamma_2 = 1$
Eigenvectors	All nonzero vectors	$u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ $u_2 = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$	$u_1 = \begin{bmatrix} 1 \\ -\mathbf{i} \end{bmatrix}$ $u_2 = \begin{bmatrix} 1 \\ +\mathbf{i} \end{bmatrix}$	$u_1 = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$	$u_1 = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$ $u_2 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}$.

Özdeğerler ve özvektörler



C. elegans sinir ağı

$$A \in \mathbb{R}^{277 \times 277}$$



$$A_{sym} := A + A^\top$$

Özdeğerler ve özvektörler

Teorem: Bir $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ kare matrisinin n tane birbirinden farklı $\lambda_1, \ldots, \lambda_n$ özdeğeri varsa, $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ özvektörleri birbirinden doğrusal bağımsızdır.

Bu vektörler $\mathbb{R}^n$ için bir baz oluşturur.

Kare bir $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ matrisinin n doğrusal bağımsız özvektörü yoksa, matris **arızalı** olarak tanımlanır. *Arızalı olmayan matrislerin* n farklı özdeğere ihtiyacı yoktur ama özvektörlerin $\mathbb{R}^n$ için bir baz oluşturması gereklidir.

$$J = \begin{bmatrix} \lambda & 1 & & \\ & \lambda & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{bmatrix} \qquad v = \begin{bmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix}$$

Özdeğerler ve özvektörler

Teorem: Bir $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ kare matrisinin n tane birbirinden farklı $\lambda_1, \ldots, \lambda_n$ özdeğerleri varsa, $\boldsymbol{x}_1, \ldots, \boldsymbol{x}_n$ özvektörleri birbirinden doğrusal bağımsızdır.

Bu vektörler $\mathbb{R}^n$ için bir baz oluşturur.

Kare bir $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ matrisinin n doğrusal bağımsız özvektörü yoksa, matris **arızalı** olarak tanımlanır. *Arızalı olmayan matrislerin* n farklı özdeğere ihtiyacı yoktur ama özvektörlerin $\mathbb{R}^n$ için bir baz oluşturması gereklidir.

Dolayısıyla arızalı matrislerin özuzaylarının boyut toplamı n'den küçüktür. Özetle, arızalı bir matrisin aritmetik katlılığı $m > 1$ olan ama geometrik katlılığı m'den küçük bir λ_i (arızalı) özdeğeri vardır.

Özdeğerler ve özvektörler

Teorem: Bir $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ matrisi için $\boldsymbol{S} := \boldsymbol{A}^\top \boldsymbol{A}$ simetrik ve yarı pozitif tanımlı bir matristir. Matrisin mertebesi $\mathrm{rk}(\boldsymbol{A}) = n$ ise $\boldsymbol{S} \in \mathbb{R}^{n \times n}$ pozitif tanımlı bir matristir.

Simetri: $\boldsymbol{S} = \boldsymbol{A}^\top \boldsymbol{A} = \boldsymbol{A}^\top (\boldsymbol{A}^\top)^\top = (\boldsymbol{A}^\top \boldsymbol{A})^\top = \boldsymbol{S}^\top$

Yarı pozitif tanımlılık: $\boldsymbol{x}^\top \boldsymbol{S} \boldsymbol{x} = \boldsymbol{x}^\top \boldsymbol{A}^\top \boldsymbol{A} \boldsymbol{x} = (\boldsymbol{x}^\top \boldsymbol{A}^\top)(\boldsymbol{A}\boldsymbol{x}) = (\boldsymbol{A}\boldsymbol{x})^\top (\boldsymbol{A}\boldsymbol{x}) \geqslant 0$

Özdeğerler ve özvektörler

Teorem: Eğer $\boldsymbol{A} \in \mathbb{R}^{n \times n}$ simetrik bir matris ise

- matrisin özdeğerleri birbirinden farklı ve reeldir.
- farklı özdeğerlere tekabül eden özvektörler birbirine ortogonaldir.
- **P** ortogonal bir matris, **D** ise köşegensel bir matris olmak üzere

$$\mathbf{A} = \mathbf{P}\mathbf{D}\mathbf{P}^{\top}$$

eşitliğini sağlayacak **P** ve **D** matrisleri bulunabilir.

Burada **D**'nin köşegenindeki değerler **A** matrisinin özdeğerleri, **P** matrisinin sütunları ise **A**'nın özvektörleridir.

Özdeğerler ve özvektörler

Örnek: $$\boldsymbol{A} = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix}$$

$$((3-\lambda)^3 + 16) - (12 + (3-\lambda))$$
$$-\lambda^3 + 9\lambda^2 - 27\lambda + 27 + 16 - 36 + 12\lambda$$
$$-\lambda^3 + 9\lambda^2 - 15\lambda + 7$$

$$p_{\boldsymbol{A}}(\lambda) = -(\lambda-1)^2(\lambda-7)$$

$$E_1 = \text{span}[\underbrace{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}}_{=:\boldsymbol{x}_1}, \underbrace{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}}_{=:\boldsymbol{x}_2}], \quad E_7 = \text{span}[\underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{=:\boldsymbol{x}_3}]$$

$$\begin{bmatrix} 2 & 2 & 2 \\ 2 & 2 & 2 \\ 2 & 2 & 2 \end{bmatrix} \mathbf{x} = \mathbf{0}$$

$$\mathbf{x} = [\, x_1 \quad x_2 \quad x_3 \,]^T$$
$$x_1 + x_2 + x_3 = 0$$
$$x_1 = -\alpha - \beta, x_2 = \alpha, x_3 = \beta,$$

$$\begin{bmatrix} -4 & 2 & 2 \\ 2 & -4 & 2 \\ 2 & 2 & -4 \end{bmatrix} \mathbf{x} = \mathbf{0}$$

$$\mathbf{x} = [\, x_1 \quad x_2 \quad x_3 \,]^T$$
$$\rightarrow \begin{bmatrix} -4 & 2 & 2 \\ 0 & -3 & 3 \\ 0 & 3 & -3 \end{bmatrix} \rightarrow \begin{bmatrix} -4 & 2 & 2 \\ 0 & -3 & 3 \\ 0 & 0 & 0 \end{bmatrix}$$
$$x_1 = x_2 = x_3 = \alpha$$

Özdeğerler ve özvektörler

Örnek: $$\boldsymbol{A} = \begin{bmatrix} 3 & 2 & 2 \\ 2 & 3 & 2 \\ 2 & 2 & 3 \end{bmatrix}$$

$$p_{\boldsymbol{A}}(\lambda) = -(\lambda - 1)^2(\lambda - 7)$$

$$E_1 = \text{span}[\underbrace{\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}}_{=:\boldsymbol{x}_1}, \underbrace{\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix}}_{=:\boldsymbol{x}_2}], \quad E_7 = \text{span}[\underbrace{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}}_{=:\boldsymbol{x}_3}]$$

$$\boldsymbol{A}(\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2) = \boldsymbol{A}\boldsymbol{x}_1 \alpha + \boldsymbol{A}\boldsymbol{x}_2 \beta = \lambda(\alpha \boldsymbol{x}_1 + \beta \boldsymbol{x}_2)$$

$$\boldsymbol{x}_1' = \begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}, \quad \boldsymbol{x}_2' = \frac{1}{2}\begin{bmatrix} -1 \\ -1 \\ 2 \end{bmatrix}$$

Gram-Schmidt

$$\mathbf{v}_1 = \mathbf{x}_1,$$
$$\mathbf{v}_2 = \mathbf{x}_2 - \frac{\langle \mathbf{x}_2, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1,$$
$$\mathbf{v}_3 = \mathbf{x}_3 - \frac{\langle \mathbf{x}_3, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 - \frac{\langle \mathbf{x}_3, \mathbf{v}_2 \rangle}{\langle \mathbf{v}_2, \mathbf{v}_2 \rangle}\mathbf{v}_2,$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\mathbf{v}_n = \mathbf{x}_n - \frac{\langle \mathbf{x}_n, \mathbf{v}_1 \rangle}{\langle \mathbf{v}_1, \mathbf{v}_1 \rangle}\mathbf{v}_1 - \cdots - \frac{\langle \mathbf{x}_n, \mathbf{v}_{n-1} \rangle}{\langle \mathbf{v}_{n-1}, \mathbf{v}_{n-1} \rangle}\mathbf{v}_{n-1}.$$

$$x_2' = x_2 - \frac{<x_1,x_2>}{<x1,x1>}x_1$$

$$\begin{bmatrix} -1 \\ 0 \\ 1 \end{bmatrix} - \frac{1}{2}\begin{bmatrix} -1 \\ 1 \\ 0 \end{bmatrix}$$

Özdeğerler ve özvektörler

$$p_{\boldsymbol{A}}(\lambda) := \det(\boldsymbol{A} - \lambda \boldsymbol{I})$$
$$= c_0 + c_1\lambda + c_2\lambda^2 + \cdots + c_{n-1}\lambda^{n-1} + (-1)^n\lambda^n$$

Teorem: $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ matrisinin determinantı (tekrarlanabilen) özdeğerlerin çarpımına eşittir. Matrisin özdeğerleri $\lambda_i \in \mathbb{C}$ ise, determinant

$$\det(\boldsymbol{A}) = \prod_{i=1}^{n} \lambda_i$$

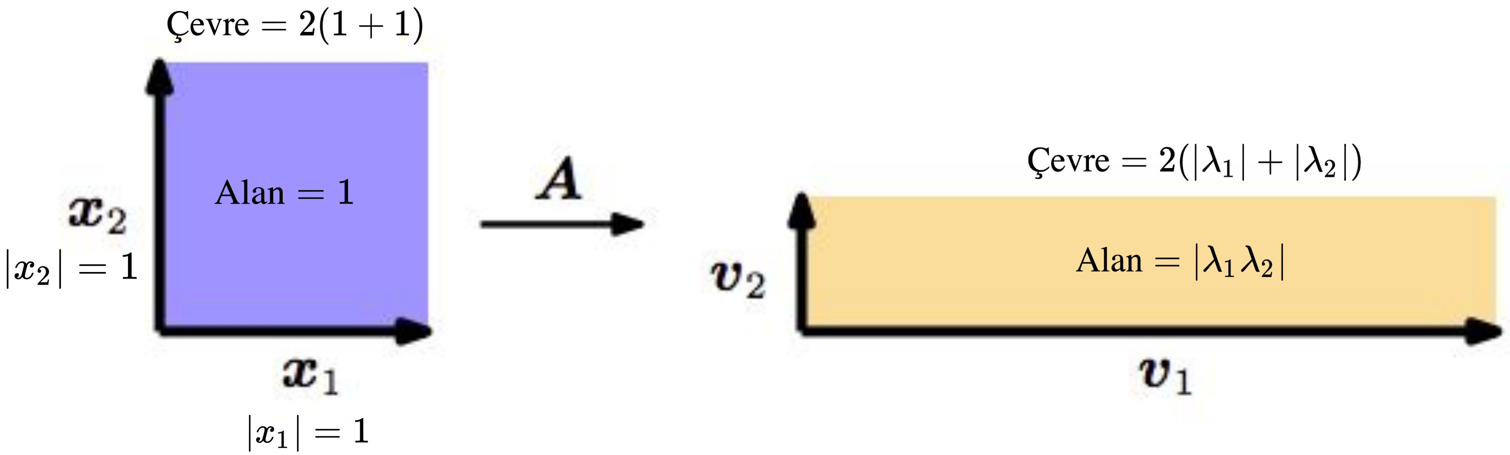
olarak hesaplanır.

Teorem: $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ matrisinin izi (tekrarlanabilen) özdeğerlerin toplamına eşittir. Matrisin özdeğerleri $\lambda_i \in \mathbb{C}$ ise, iz

$$tr(\boldsymbol{A}) = \sum_{i=1}^{n} \lambda_i$$

olarak hesaplanır.

Özdeğerler ve özvektörler



$$\boldsymbol{v}_1 = \boldsymbol{A}\boldsymbol{x}_1 = \lambda_1 \boldsymbol{x}_1$$

$$\boldsymbol{v}_2 = \boldsymbol{A}\boldsymbol{x}_2 = \lambda_2 \boldsymbol{x}_2$$

Cholesky ayrıştırması

Teorem: Simetrik ve pozitif kesin bir $\boldsymbol{A}$ matrisi, $\boldsymbol{L}$ köşegen değerleri pozitif olan alt üçgensel bir matris olmak üzere $\boldsymbol{A} = \boldsymbol{L}\boldsymbol{L}^\top$ şeklinde çarpanlarına ayrılabilir.

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} l_{11} & \cdots & 0 \\ \vdots & \ddots & \vdots \\ l_{n1} & \cdots & l_{nn} \end{bmatrix} \begin{bmatrix} l_{11} & \cdots & l_{n1} \\ \vdots & \ddots & \vdots \\ 0 & \cdots & l_{nn} \end{bmatrix}$$

Bu ayrışmayı sağlayan sadece tek bir $\boldsymbol{L}$ matrisi vardır ve $\boldsymbol{A}$ matrisinin **Cholesky çarpanı** olarak tanımlanır.

Cholesky ayrıştırması

Örnek: 3x3 bir simetrik, pozitif kesin matris aşağıdaki şekilde Cholesky çarpanlarına ayrılır:

$$\boldsymbol{A} = \begin{bmatrix} a_{11} & a_{21} & a_{31} \\ a_{21} & a_{22} & a_{32} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = \boldsymbol{L}\boldsymbol{L}^{\top} = \begin{bmatrix} l_{11} & 0 & 0 \\ l_{21} & l_{22} & 0 \\ l_{31} & l_{32} & l_{33} \end{bmatrix} \begin{bmatrix} l_{11} & l_{21} & l_{31} \\ 0 & l_{22} & l_{32} \\ 0 & 0 & l_{33} \end{bmatrix}$$

$$\boldsymbol{A} = \begin{bmatrix} l_{11}^2 & l_{21}l_{11} & l_{31}l_{11} \\ l_{21}l_{11} & l_{21}^2 + l_{22}^2 & l_{31}l_{21} + l_{32}l_{22} \\ l_{31}l_{11} & l_{31}l_{21} + l_{32}l_{22} & l_{31}^2 + l_{32}^2 + l_{33}^2 \end{bmatrix}$$

$$l_{11} = \sqrt{a_{11}}\,, \quad l_{22} = \sqrt{a_{22} - l_{21}^2}\,, \quad l_{33} = \sqrt{a_{33} - (l_{31}^2 + l_{32}^2)}$$

$$l_{21} = \frac{1}{l_{11}} a_{21}\,, \quad l_{31} = \frac{1}{l_{11}} a_{31}\,, \quad l_{32} = \frac{1}{l_{22}} (a_{32} - l_{31}l_{21})$$

Cholesky ayrıştırması

- Cholesky ayrıştırması doğrusal sistem çözümlerinde büyük kolaylık sağlar.
    - Klasik LU ayrıştırmasının yarısı kadar hafıza harcar.
    - Yine $Ax = b$ çözümü için göre yaklaşık iki kat daha hızlı sonuç verir.
- Monte Carlo algoritmalarında belli bir korelasyona sahip rassal değerlerin oluşturulması esnasında kovaryans matrisinin çarpanlarına ayrılması için kullanılır.
- Daha fazla uygulama için Wikipedia’ya bakabilirsiniz.

Özayrıştırma ve köşegenleştirme

Koşegen (diyagonal) matrislerin determinant, üs alma, tersini bulma gibi hesaplamaları diğer matrislere göre oldukça kolaydır.

$$\boldsymbol{D} = \begin{bmatrix} c_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & c_n \end{bmatrix}$$

- Determinant, köşegendeki değerlerin çarpımına eşittir.
- Matrisin *k*'nıncı kuvveti, elemanların *k*'nıncı kuvvetini alarak elde edilir.
- Matrisin tersi her elemanın çarpmaya göre tersi alınarak elde edilir.

Özayrıştırma ve köşegenleştirme

Tanım: $\boldsymbol{A}, \tilde{\boldsymbol{A}} \in \mathbb{R}^{n \times n}$ matrisleri $\tilde{\boldsymbol{A}} = \boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$ eşitliğini sağlayan düzenli (tersi alınabilen) bir $\boldsymbol{S} \in \mathbb{R}^{n \times n}$ matrisi bulunabiliyorsa birbirine **benzerdir**.

Özayrıştırma için üzerinde duracağımız matrisler, köşegen matrislere benzer matrislerdir. Özetle ilgilendiğimiz ayrıştırma

$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}$$

şeklindeki ayrıştırmalardır.

Özayrıştırma ve köşegenleştirme

Tanım: Bir $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ matrisi için $\boldsymbol{D} = \boldsymbol{P}^{-1}\boldsymbol{A}\boldsymbol{P}$ köşegen matrisi olacak şekilde bir $\boldsymbol{P} \in \mathbb{R}^{n\times n}$ matrisi bulunabiliyorsa, $\boldsymbol{A}$ matrisi köşegen $\boldsymbol{D}$ matrisi ile benzerdir veya **köşegenleştirilebilir** bir matristir.

$$\boldsymbol{D} = \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} \qquad \boldsymbol{P} := [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n]$$

olsun.

$$\boldsymbol{A}\boldsymbol{P} = \boldsymbol{P}\boldsymbol{D}$$

eşitliğinin ancak ve ancak $\lambda_1, \ldots, \lambda_n$ değerleri $\boldsymbol{A}$ 'nın özdeğerleri, $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n$ de $\boldsymbol{A}$ 'nın özvektörleri ise sağlanacağını şu şekilde gösterebiliriz.

Özayrıştırma ve köşegenleştirme

$$\boldsymbol{AP} = \boldsymbol{A}[\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n] = [\boldsymbol{A}\boldsymbol{p}_1, \ldots, \boldsymbol{A}\boldsymbol{p}_n],$$

$$\boldsymbol{PD} = [\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n] \begin{bmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{bmatrix} = [\lambda_1 \boldsymbol{p}_1, \ldots, \lambda_n \boldsymbol{p}_n]$$

Burada $\boldsymbol{P} \in \mathbb{R}^{n \times n}$ matrisi tersi alınabilir bir matris olduğundan, mertebesi n olmalıdır. Bu da $\boldsymbol{p}_1, \ldots, \boldsymbol{p}_n$ vektörlerinin birbirinden doğrusal bağımsız ve $\mathbb{R}^n$ için bir baz olmasını gerektirir.

Özayrıştırma ve köşegenleştirme

Teorem: Bir $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ matrisi $\boldsymbol{P} \in \mathbb{R}^{n\times n}$ ve $\boldsymbol{D}$ matrisi $\boldsymbol{A}$'nın özdeğerlerini içeren köşegen bir matris olmak üzere

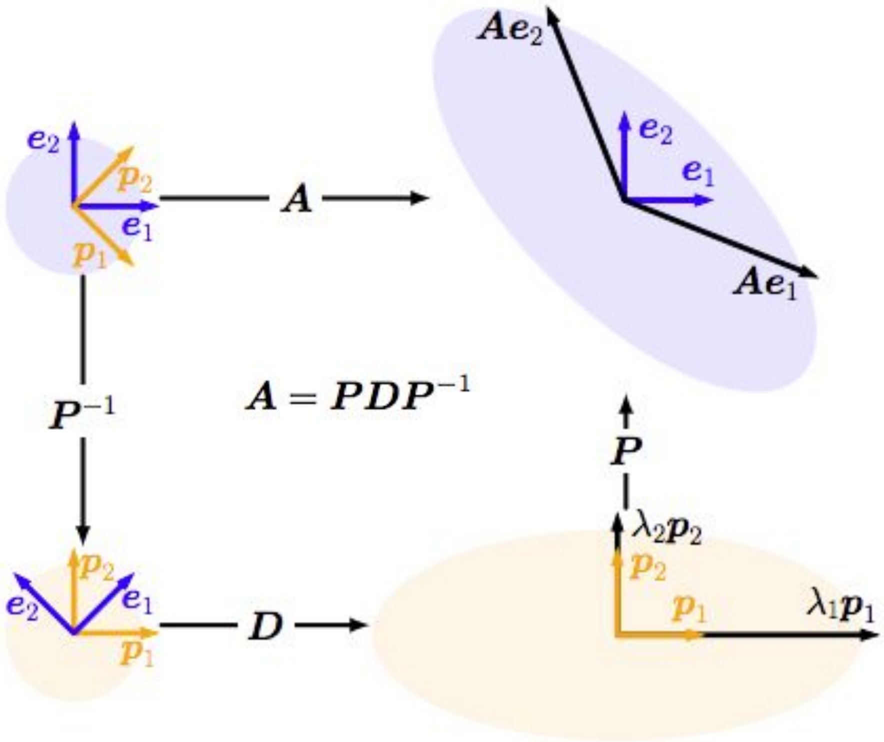
$$\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}$$

şeklinde çarpanlarına ancak ve ancak özvektörleri $\mathbb{R}^n$ için bir baz oluşturuyorsa ayrılabilir.

Dolayısıyla sadece arızalı olmayan matrislerin özayrıştılması mümkündür ve $\boldsymbol{P} \in \mathbb{R}^{n\times n}$ *matrisinin sütunları* $\boldsymbol{A}$*'nın özvektörleridir.*

Teorem: Bir simetrik matris her zaman köşegenleştirilebilir. Burada $\boldsymbol{P} \in \mathbb{R}^{n\times n}$ ortagonal bir matristir (spektral teorem).

Özayrıştırma ve köşegenleştirme



Özayrıştırma ve köşegenleştirme

Örnek: $\boldsymbol{A} = \frac{1}{2}\begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}$ matrisinin özayrıştırmasını bulunuz.

$$\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = \det\left(\begin{bmatrix} \frac{5}{2} - \lambda & -1 \\ -1 & \frac{5}{2} - \lambda \end{bmatrix}\right)$$
$$= (\tfrac{5}{2} - \lambda)^2 - 1 = \lambda^2 - 5\lambda + \tfrac{21}{4} = (\lambda - \tfrac{7}{2})(\lambda - \tfrac{3}{2})$$

$$\boldsymbol{\lambda_1 = \tfrac{7}{2}}$$
$$\boldsymbol{\lambda_2 = \tfrac{3}{2}}$$

$$\frac{1}{2}\begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}\mathbf{p}_1 = \frac{7}{2}\mathbf{p}_1 \qquad \begin{aligned} 5\alpha - 2\beta &= 7\alpha \\ -2\alpha + 5\beta &= 7\beta \end{aligned} \qquad \alpha = -\beta \qquad \mathbf{p}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\frac{1}{2}\begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}\mathbf{p}_2 = \frac{3}{2}\mathbf{p}_2 \qquad \begin{aligned} 5\alpha - 2\beta &= 3\alpha \\ -2\alpha + 5\beta &= 3\beta \end{aligned} \qquad \alpha = \beta \qquad \mathbf{p}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

Bu vektörler $\mathbb{R}^2$ için bir baz oluşturduğundan matrisin özayrıştırması vardır.

Özayrıştırma ve köşegenleştirme

Örnek: $\boldsymbol{A} = \frac{1}{2}\begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}$ matrisinin özayrıştırmasını bulunuz.

$$\det(\boldsymbol{A} - \lambda \boldsymbol{I}) = \det\left(\begin{bmatrix} \frac{5}{2} - \lambda & -1 \\ -1 & \frac{5}{2} - \lambda \end{bmatrix}\right)$$
$$= (\tfrac{5}{2} - \lambda)^2 - 1 = \lambda^2 - 5\lambda + \tfrac{21}{4} = (\lambda - \tfrac{7}{2})(\lambda - \tfrac{3}{2})$$

$$\lambda_1 = \frac{7}{2}$$

$$\lambda_2 = \frac{3}{2}$$

$$\mathbf{p}_1 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ -1 \end{bmatrix}$$

$$\mathbf{p}_2 = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

$$\boldsymbol{P} = [\boldsymbol{p}_1,\ \boldsymbol{p}_2] = \frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}$$

$$\underbrace{\frac{1}{2}\begin{bmatrix} 5 & -2 \\ -2 & 5 \end{bmatrix}}_{\boldsymbol{A}} = \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}}_{\boldsymbol{P}} \underbrace{\begin{bmatrix} \frac{7}{2} & 0 \\ 0 & \frac{3}{2} \end{bmatrix}}_{\boldsymbol{D}} \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}}_{\boldsymbol{P}^{-1}}$$

Özayrıştırma ve köşegenleştirme

Matrisin $\boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}$ şeklinde bir özayrıştırması varsa

- üs alma işlemi kolaylıkla yapılabilir

$$\boldsymbol{A}^k = (\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1})^k = \boldsymbol{P}\boldsymbol{D}^k\boldsymbol{P}^{-1}$$

- determinant hesabı kolaylıkla yapılabilir

$$\begin{aligned}\det(\boldsymbol{A}) = \det(\boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^{-1}) &= \det(\boldsymbol{P})\det(\boldsymbol{D})\det(\boldsymbol{P}^{-1}) \\ &= \det(\boldsymbol{D}) = \prod_i d_{ii}\end{aligned}$$

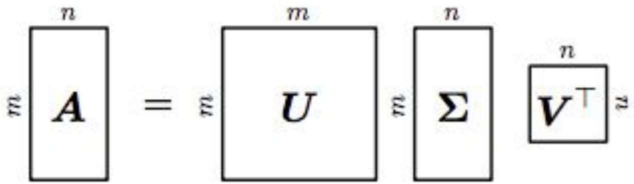
Tekil değer ayrıştırması

Tekil değer ayrıştırması (SVD) sadece kare, simetrik ya da arızasızlara değil, bütün matrislere, her zaman uygulanabilir.

Eğer $\mathbf{A} \in \mathbb{R}^{m \times n}$ doğrusal bir eşleme $\Phi : V \to W$ ise SVD ilgili vektör uzaylarının geometrisi arasındaki değişimin değerini hesaplayabilir.

Tekil değer ayrıştırması

Teorem: $\mathbf{A} \in \mathbb{R}^{m \times n}$ mertebesi $r \in [0, min(m,n)]$ olan dikdörtgen bir matris olsun. $\mathbf{A}$'nın SVDsi



şeklinde hesaplanır. Burada $\mathbf{U} \in \mathbb{R}^{m \times m}$ sütunları $\mathbf{u}_i, \; i = 1, \ldots, m$ olan ortogonal bir matris, $\mathbf{V} \in \mathbb{R}^{n \times n}$ sütunları $\mathbf{v}_j, \; j = 1, \ldots, n$ olan ortogonal bir matristir. $\mathbf{\Sigma}$ ise değerleri $\mathbf{\Sigma}_{ij} = 0, \; i \neq j$ ve $\mathbf{\Sigma}_{ii} = \sigma_i > 0$ özelliklerini sağlayan $m \times n$ boyutlarında bir matristir.

Özayrıştırma ile SVD'yi karşılaştıracak olursak

$$\boldsymbol{S} = \boldsymbol{S}^\top = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^\top \qquad \boldsymbol{U} = \boldsymbol{P} = \boldsymbol{V}, \quad \boldsymbol{D} = \boldsymbol{\Sigma}$$

$$\boldsymbol{S} = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top$$

Tekil değer ayrıştırması

$\mathbf{\Sigma}$ matrisinin köşegeninde bulunan $\sigma_i, \; i = 1, \ldots, r$ değerlerine $\mathbf{A}$ matrisinin tekil değerleri denir. $\mathbf{u}_i$ vektörleri matrisin sol tekil vektörleri, $\mathbf{v}_j$ vektörleri ise sağ tekil vektörleri olarak adlandırılır.

$\sigma_1 \geq \sigma_2 \geq \ldots \geq \sigma_r$ olarak kabul edilir.

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_n \\ 0 & \ldots & 0 \\ \vdots & & \vdots \\ 0 & \ldots & 0 \end{bmatrix}$$

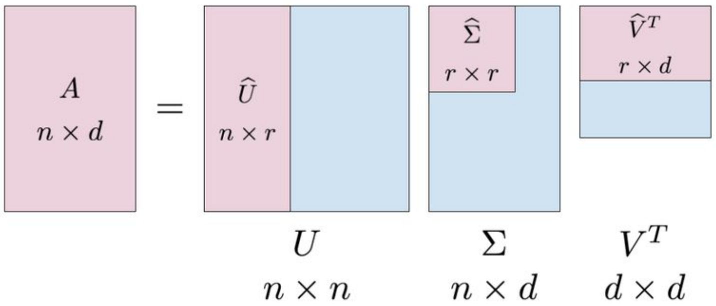
$$m > n$$

$$\mathbf{\Sigma} = \begin{bmatrix} \sigma_1 & 0 & 0 & 0 & \ldots & 0 \\ 0 & \ddots & 0 & \vdots & & \vdots \\ 0 & 0 & \sigma_m & 0 & \ldots & 0 \end{bmatrix}$$

$$m < n$$

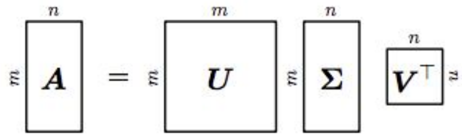
Tekil değer ayrıştırması

$\mathbf{\Sigma}$ matrisinin köşegeninde bulunan $\sigma_i,\ i = 1, \ldots, r$ değerlerine $\mathbf{A}$ matrisinin tekil değerleri denir. $\mathbf{u}_i$ vektörleri matrisin sol tekil vektörleri, $\mathbf{v}_j$ vektörleri ise sağ tekil vektörleri olarak adlandırılır.



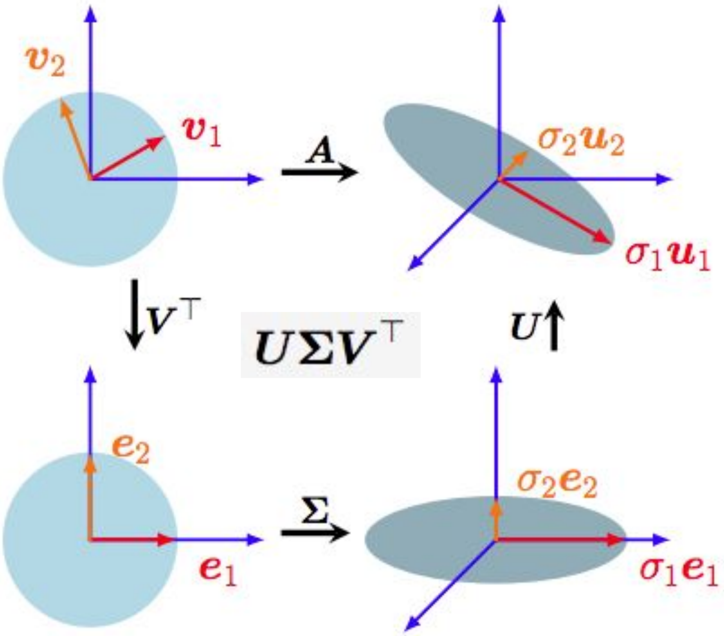
$$\mathbf{A} = \mathbf{U\Sigma V}^T \rightarrow \mathbf{AV} = \mathbf{U\Sigma} \rightarrow \begin{bmatrix} \mathbf{Av}_1 & = \sigma_1 \mathbf{u}_1 \\ \mathbf{Av}_2 & = \sigma_2 \mathbf{u}_2 \\ & \cdots \\ \mathbf{Av}_r & = \sigma_r \mathbf{u}_r \end{bmatrix}$$

Tekil değer ayrıştırması

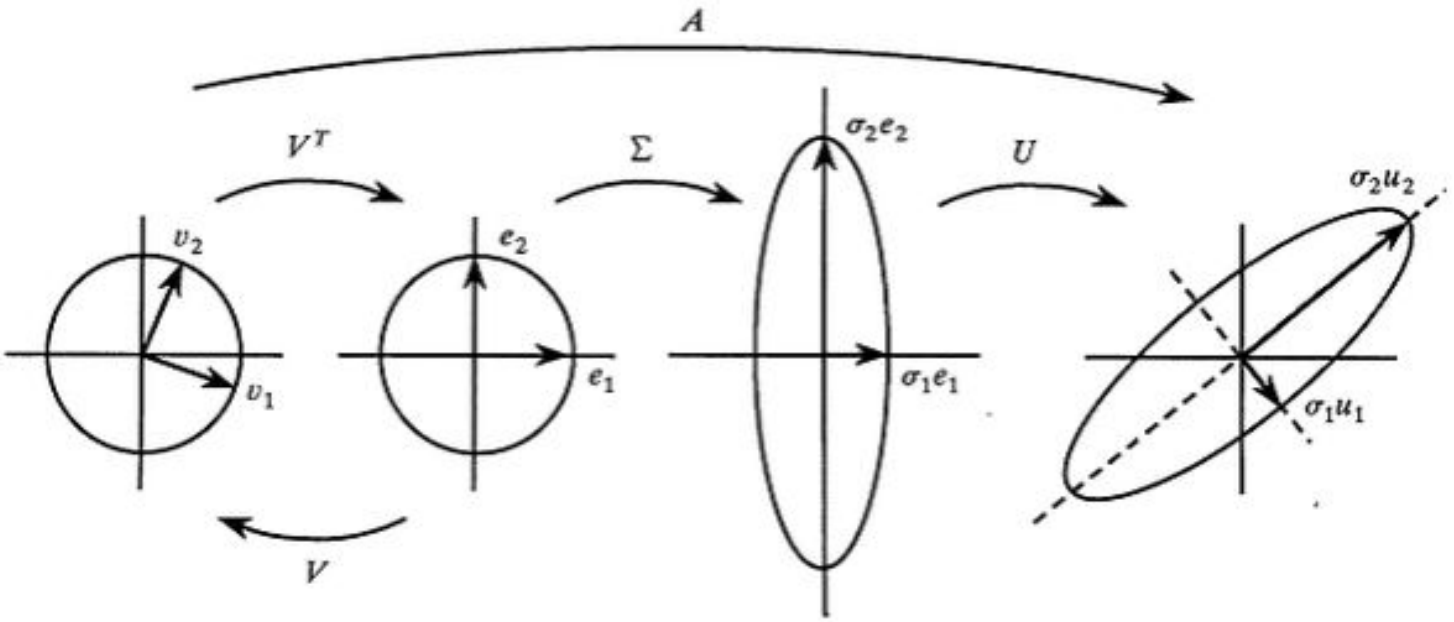


SVD ayıştırması $\mathbf{A}$ matrisinin uyguladığı dönüşümü 3'e ayırır.

1. Önce $\mathbf{V}^T$ ile $\mathbb{R}^n$ uzayında gerekli baz değişikliği yapılır. Bu sayede yeni baz $\mathbf{v}_j,\ j = 1, \ldots, n$ vektörleri olur.
2. Daha sonra $\mathbf{\Sigma}$ ile esnetme/kısaltma işlemi gerçekleşir. Burada her baz vektörü ilgili tekil değer ile çarpılır. Bunun yanında yeni boyut eklenebilir (ya da $m < n$ ise çıkarılabilir).
3. Son olarak $\mathbf{U}$ matrisi ile geometri tekrar standart baza dönüştürülür.



Tekil değer ayrıştırması

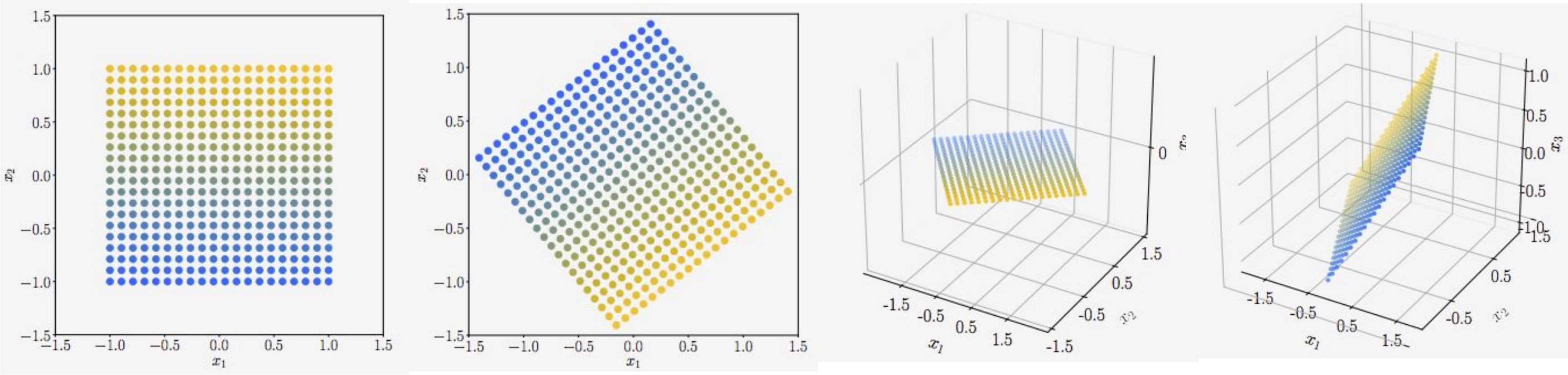


Tekil değer ayrıştırması

Örnek:

$$
\boldsymbol{A} = \begin{bmatrix} 1 & -0.8 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} = \boldsymbol{U\Sigma V}^{\top}
$$

$$
= \begin{bmatrix} -0.79 & 0 & -0.62 \\ 0.38 & -0.78 & -0.49 \\ -0.48 & -0.62 & 0.62 \end{bmatrix} \begin{bmatrix} 1.62 & 0 \\ 0 & 1.0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} -0.78 & 0.62 \\ -0.62 & -0.78 \end{bmatrix}
$$



Tekil değer ayrıştırması

SVD için $\mathbf{U}$ ve $\mathbf{V}$ matrislerinin bulunmasına ihtiyaç vardır. Önce $\mathbf{V}$ matrisinden başlayalım.

$$\boldsymbol{A}^\top \boldsymbol{A} = \boldsymbol{P}\boldsymbol{D}\boldsymbol{P}^\top = \boldsymbol{P}\begin{bmatrix} \lambda_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \lambda_n \end{bmatrix}\boldsymbol{P}^\top$$

eşitliğinde $\boldsymbol{A}^\top \boldsymbol{A}$ simetrik olduğundan eşitliği sağlayan matrisler bulunabilir. Eğer $\mathbf{A}$ matrisinin SVDsi var ise $\boldsymbol{A}^\top \boldsymbol{A}$ şu şekilde de yazılabilir.

$$\boldsymbol{A}^\top \boldsymbol{A} = (\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top)^\top(\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top) = \boldsymbol{V}\boldsymbol{\Sigma}^\top\boldsymbol{U}^\top\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top$$

$\mathbf{U}$ ortonormal bir matris olduğundan $\boldsymbol{U}^\top \boldsymbol{U} = \boldsymbol{I}$ olur

$$\boldsymbol{A}^\top \boldsymbol{A} = \boldsymbol{V}\boldsymbol{\Sigma}^\top\boldsymbol{\Sigma}\boldsymbol{V}^\top = \boldsymbol{V}\begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_n^2 \end{bmatrix}\boldsymbol{V}^\top$$

$$\boldsymbol{V}^\top = \boldsymbol{P}^\top$$
$$\sigma_i^2 = \lambda_i$$

Tekil değer ayrıştırması

$\mathbf{U}$ matrisinin hesabı benzerdir.

$$\boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{S}\boldsymbol{D}\boldsymbol{S}^\top$$

eşitliğinde $\boldsymbol{A}\boldsymbol{A}^\top$ simetrik bir matris olduğundan bunu sağlayan matrisler bulunabilir. Eğer $\mathbf{A}$ matrisinin SVDsi var ise $\boldsymbol{A}\boldsymbol{A}^\top$ şu şekilde de yazılabilir.

$$\boldsymbol{A}\boldsymbol{A}^\top = (\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top)(\boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top)^\top = \boldsymbol{U}\boldsymbol{\Sigma}\boldsymbol{V}^\top\boldsymbol{V}\boldsymbol{\Sigma}^\top\boldsymbol{U}^\top$$

$\mathbf{V}$ ortonormal bir matris olduğundan $\mathbf{V}^T\mathbf{V} = \mathbf{I}$ olur ve

$$\boldsymbol{A}\boldsymbol{A}^\top = \boldsymbol{U}\begin{bmatrix} \sigma_1^2 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \sigma_m^2 \end{bmatrix}\boldsymbol{U}^\top$$

$\boldsymbol{A}\boldsymbol{A}^\top$ ve $\boldsymbol{A}^\top\boldsymbol{A}$ ‘nın özdeğerleri aynı olduğundan $\boldsymbol{\Sigma}$ matrisleri de aynıdır.

Tekil değer ayrıştırma

Aslında bir önceki yansıdaki, **U** için yaptığımız işlemlere ihtiyacımız yok.

$$\mathbf{A} = \mathbf{U\Sigma V}^T \rightarrow \mathbf{AV} = \mathbf{U\Sigma} \longrightarrow \begin{aligned} \mathbf{Av}_1 &= \sigma_1 \mathbf{u}_1 \\ \mathbf{Av}_2 &= \sigma_2 \mathbf{u}_2 \\ &\dots \\ \mathbf{Av}_r &= \sigma_r \mathbf{u}_r \end{aligned} \longrightarrow \mathbf{u}_i = \frac{\mathbf{Av}_i}{\sigma_i}, 1 \leq i \leq r$$

Burada $\mathbf{u}_i \ , 1 \leq i \leq r$ ortonormal bir baz oluşturuyor:

$$(\boldsymbol{A}\boldsymbol{v}_i)^\top(\boldsymbol{A}\boldsymbol{v}_j) = \boldsymbol{v}_i^\top(\boldsymbol{A}^\top\boldsymbol{A})\boldsymbol{v}_j = \boldsymbol{v}_i^\top(\lambda_j\boldsymbol{v}_j) = \lambda_j\boldsymbol{v}_i^\top\boldsymbol{v}_j = 0$$

$$\mathbf{u}_i^T\mathbf{u}_i = \frac{\lambda_i \mathbf{v}_i^T \mathbf{v}_i}{\sigma_i{}^2} = 1$$

Tekil değer ayrıştırması

Örnek:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix}$$

$$\boldsymbol{A}^\top \boldsymbol{A} = \begin{bmatrix} 1 & -2 \\ 0 & 1 \\ 1 & 0 \end{bmatrix} \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} = \begin{bmatrix} 5 & -2 & 1 \\ -2 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix}$$

$$det(\begin{pmatrix} 5-x & -2 & 1 \\ -2 & 1-x & 0 \\ 1 & 0 & 1-x \end{pmatrix})$$

$$-x^3 + 7x^2 - 6x = 0$$

$$\boldsymbol{A}^\top \boldsymbol{A} = \begin{bmatrix} \frac{5}{\sqrt{30}} & 0 & \frac{-1}{\sqrt{6}} \\ \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}} \end{bmatrix} \begin{bmatrix} 6 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \frac{5}{\sqrt{30}} & \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{30}} \\ 0 & \frac{1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{-1}{\sqrt{6}} & \frac{-2}{\sqrt{6}} & \frac{1}{\sqrt{6}} \end{bmatrix} = \boldsymbol{PDP}^\top$$

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sqrt{6} & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$\boldsymbol{V} = \boldsymbol{P} = \begin{bmatrix} \frac{5}{\sqrt{30}} & 0 & \frac{-1}{\sqrt{6}} \\ \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

Tekil değer ayrıştırması

Örnek:

$$\boldsymbol{A} = \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix}$$

$$\boldsymbol{V} = \boldsymbol{P} = \begin{bmatrix} \frac{5}{\sqrt{30}} & 0 & \frac{-1}{\sqrt{6}} \\ \frac{-2}{\sqrt{30}} & \frac{1}{\sqrt{5}} & \frac{-2}{\sqrt{6}} \\ \frac{1}{\sqrt{30}} & \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{6}} \end{bmatrix}$$

$$\boldsymbol{u}_1 = \frac{1}{\sigma_1}\boldsymbol{A}\boldsymbol{v}_1 = \frac{1}{\sqrt{6}} \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} \frac{5}{\sqrt{30}} \\ \frac{-2}{\sqrt{30}} \\ \frac{1}{\sqrt{30}} \end{bmatrix} = \begin{bmatrix} \frac{1}{\sqrt{5}} \\ -\frac{2}{\sqrt{5}} \end{bmatrix}$$

$$\boldsymbol{u}_2 = \frac{1}{\sigma_2}\boldsymbol{A}\boldsymbol{v}_2 = \frac{1}{1} \begin{bmatrix} 1 & 0 & 1 \\ -2 & 1 & 0 \end{bmatrix} \begin{bmatrix} 0 \\ \frac{1}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} \end{bmatrix} = \begin{bmatrix} \frac{2}{\sqrt{5}} \\ \frac{1}{\sqrt{5}} \end{bmatrix},$$

$$\boldsymbol{U} = [\boldsymbol{u}_1, \boldsymbol{u}_2] = \frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ -2 & 1 \end{bmatrix}$$

Tekil değer ayrıştırması

- SVD ve özayrıştırmanın farkları
    - SVD bütün reel matrisler için var, öz ayrıştırma ise sadece kare matrisler için tanımlı. O da sadece eğer özvektörlerden ilgili uzay için bir baz oluşturabilirsek mümkün.
    - İki ayrıştırmada da operasyon (tanım uzayında) baz değişimi, baz vektörlerinin ayrı ayrı ölçeklendirilmesi ve (değer uzayında) baz değişimi şeklinde. Ama SVD'de tanım ve değer uzayı'nın boyutları farklı (olabilir).
    - Özayrıştırmada **P** matrisinin içerisindeki vektörler birbirine ortagonal olmak zorunda değil. Bu da baz değişimi işleminin ve **A**'nın yarattığı dönüşümün rotasyon ve ölçekleme gibi parçalara ayrılmasını engelliyor.

Tekil değer ayrıştırması

- SVD ve özayrıştırmanın farkları
  - Özayrıştırmada **P** ve **P**'nin tersi kullanılırken SVD'de **U** ve **V** matrisleri birbirinin tersi olmak zorunda değil.
  - SVD'de diagonal matris değerleri reel ve sıfırdan büyük eşit. Özayrıştırmada böyle bir garanti yok.
  - Yine de bunlar birbirleri ile alakalı ayrıştırmalar:
    - **A** matrisinin sol tekil vektörleri $\boldsymbol{A}\boldsymbol{A}^\top$ matrisinin özvektörleridir.
    - **A** matrisinin sağ tekil vektörleri $\boldsymbol{A}^\top\boldsymbol{A}$ matrisinin özvektörleridir.
    - **A** matrisinin sıfırdan farklı tekil değerleri $\boldsymbol{A}^\top\boldsymbol{A}$ matrisinin (dolayısıyla $\boldsymbol{A}\boldsymbol{A}^\top$ matrisinin) özdeğerlerinin kareköklerine eşittir.

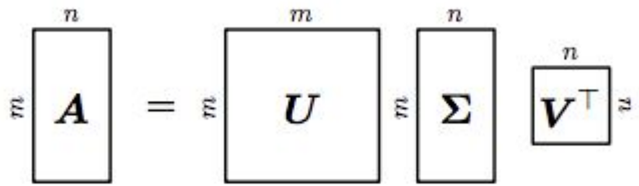
Tekil değer ayrıştırması

$$
\begin{array}{r c}
 & \begin{array}{ccc} \text{Ali} & \text{Beatrix} & \text{Chandra} \end{array} \\
\begin{array}{r} \text{Star Wars} \\ \text{Blade Runner} \\ \text{Amelie} \\ \text{Delicatessen} \end{array} &
\begin{bmatrix} 5 & 4 & 1 \\ 5 & 5 & 0 \\ 0 & 0 & 5 \\ 1 & 0 & 4 \end{bmatrix}
\end{array}
=
\begin{array}{c}
\begin{array}{cccc} \text{Bilimkurgu} & \text{Sanat Filmi} & ? & ? \end{array} \\
\begin{bmatrix} -0.6710 & 0.0236 & 0.4647 & -0.5774 \\ -0.7197 & 0.2054 & -0.4759 & 0.4619 \\ -0.0939 & -0.7705 & -0.5268 & -0.3464 \\ -0.1515 & -0.6030 & 0.5293 & -0.5774 \end{bmatrix}
\end{array}
$$

$$
\begin{bmatrix} 9.6438 & 0 & 0 \\ 0 & 6.3639 & 0 \\ 0 & 0 & 0.7056 \\ 0 & 0 & 0 \end{bmatrix}
$$

$$
\begin{array}{r c}
\begin{array}{r} \text{Bilimkurgu sever} \\ \text{Sanat filmi sever} \\ ? \end{array} &
\begin{bmatrix} -0.7367 & -0.6515 & -0.1811 \\ 0.0852 & 0.1762 & -0.9807 \\ 0.6708 & -0.7379 & -0.0743 \end{bmatrix} \\
 & \begin{array}{ccc} \text{Ali} & \text{Beatrix} & \text{Chandra} \end{array}
\end{array}
$$

Matris yaklaşımı

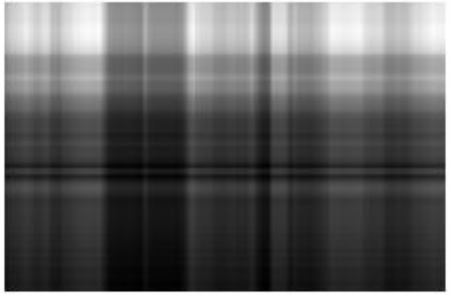


$$\boldsymbol{A} = \sum_{i=1}^{r} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^{\top} = \sum_{i=1}^{r} \sigma_i \boldsymbol{A}_i$$

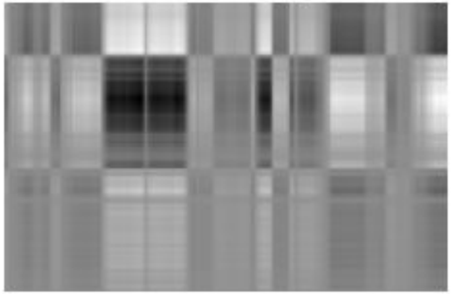
Matris yaklaşımı

$$\boldsymbol{A} \in \mathbb{R}^{1432\times1910}$$

(a) Original image $\boldsymbol{A}$.



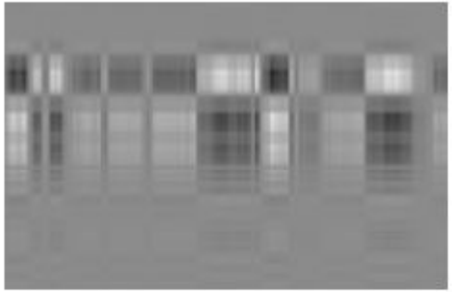
(b) $\boldsymbol{A}_1,\ \sigma_1 \approx 228,052.$



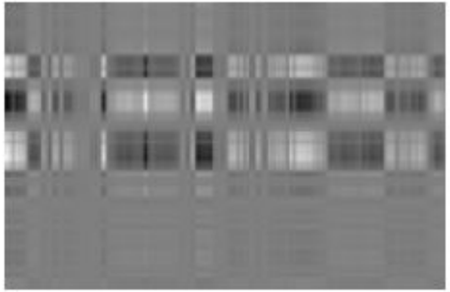
(c) $\boldsymbol{A}_2,\ \sigma_2 \approx 40,647.$

$$\boldsymbol{A}_i := \boldsymbol{u}_i \boldsymbol{v}_i^\top$$

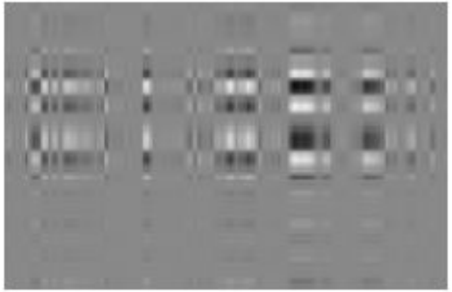
```
u = [1,2,3]'

v = [5,10,15]'

 5    10    15
10    20    30
15    30    45
```



(d) $\boldsymbol{A}_3,\ \sigma_3 \approx 26,125.$



(e) $\boldsymbol{A}_4,\ \sigma_4 \approx 20,232.$

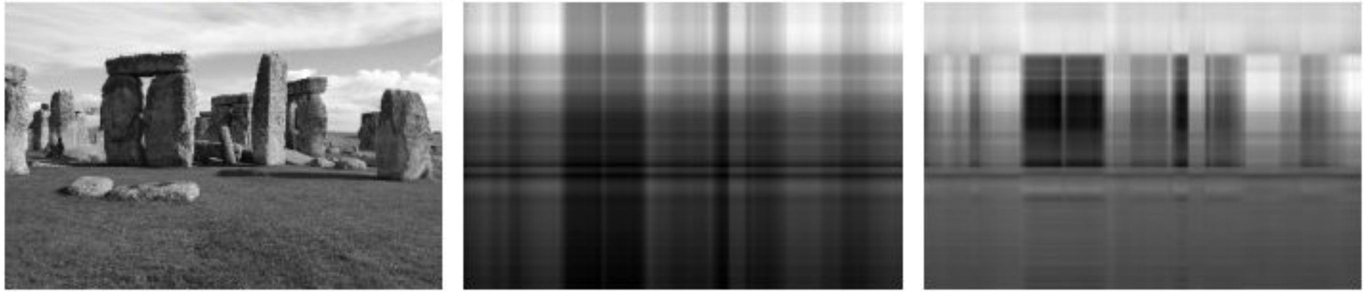


(f) $\boldsymbol{A}_5,\ \sigma_5 \approx 15,436.$

Matris yaklaşımı

$$\boldsymbol{A} \in \mathbb{R}^{1432\times 1910}$$

$$\begin{aligned}\widehat{\boldsymbol{A}}(k) &:= \sum_{i=1}^{k} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^\top \\ &= \sum_{i=1}^{k} \sigma_i \boldsymbol{A}_i\end{aligned}$$



(a) Original image $\boldsymbol{A}$. (b) Rank-1 approximation $\widehat{\boldsymbol{A}}(1)$. (c) Rank-2 approximation $\widehat{\boldsymbol{A}}(2)$.

(d) Rank-3 approximation $\widehat{\boldsymbol{A}}(3)$. (e) Rank-4 approximation $\widehat{\boldsymbol{A}}(4)$. (f) Rank-5 approximation $\widehat{\boldsymbol{A}}(5)$.

Matris yaklaşımı

$$\widehat{\boldsymbol{A}}(k) := \sum_{i=1}^{k} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^\top = \sum_{i=1}^{k} \sigma_i \boldsymbol{A}_i$$

$\boldsymbol{A} \in \mathbb{R}^{1432 \times 1910}$

(a) Original image $\boldsymbol{A}$.

(f) Rank-5 approximation $\widehat{\boldsymbol{A}}(5)$.

1432 x 1910 = 2,735,120

5 x (1432 + 1910 + 1) = 16,715
(orjinal boyutun sadece %0.6sı)

Matris yaklaşımı

Matris normları:

$$\|A\|_1 = \max_{1\le j\le n} \sum_{i=1}^{m} |a_{ij}|,$$

$$\|A\|_2 = \sigma_{\max}(A)$$

$$\|A\|_\infty = \max_{1\le i\le m} \sum_{j=1}^{n} |a_{ij}|.$$

$$\|A\|_{\mathrm{F}} = \left(\sum_{i=1}^{m} \sum_{j=1}^{n} |a_{ij}|^2 \right)^{\frac{1}{2}}$$

Matris yaklaşımı

Tanım: Bir $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ matrisinin **spektral normu** bütün $\boldsymbol{x} \in \mathbb{R}^n \setminus \{\mathbf{0}\}$ vektörleri üzerinden şu şekilde hesaplanır

$$\|\boldsymbol{A}\|_2 := \max_{\boldsymbol{x}} \frac{\|\boldsymbol{A}\boldsymbol{x}\|_2}{\|\boldsymbol{x}\|_2}$$

Teorem: Bir matrisin spektral normu en büyük tekil değerine eşittir.

Teorem (Eckart-Young): $\boldsymbol{A} \in \mathbb{R}^{m \times n}$ mertebesi *r* olan, $\boldsymbol{B} \in \mathbb{R}^{m \times n}$ ise mertebesi *k* ≤ *r*, olan iki matris olsun. Mertebesi *k* olan yaklaşık matris $\widehat{\boldsymbol{A}}(k) = \sum_{i=1}^{k} \sigma_i \boldsymbol{u}_i \boldsymbol{v}_i^\top$ için

$$\widehat{\boldsymbol{A}}(k) = \operatorname{argmin}_{\operatorname{rk}(\boldsymbol{B})=k} \|\boldsymbol{A} - \boldsymbol{B}\|_2 \,,$$

$$\left\|\boldsymbol{A} - \widehat{\boldsymbol{A}}(k)\right\|_2 = \sigma_{k+1} \,.$$

ifadesi doğrudur.

Matris yaklaşımı

$$
\begin{array}{r|ccc|}
 & \text{Ali} & \text{Beatrix} & \text{Chandra} \\
\text{Star Wars} & 5 & 4 & 1 \\
\text{Blade Runner} & 5 & 5 & 0 \\
\text{Amelie} & 0 & 0 & 5 \\
\text{Delicatessen} & 1 & 0 & 4
\end{array}
=
\begin{bmatrix}
-0.6710 & 0.0236 & 0.4647 & -0.5774 \\
-0.7197 & 0.2054 & -0.4759 & 0.4619 \\
-0.0939 & -0.7705 & -0.5268 & -0.3464 \\
-0.1515 & -0.6030 & 0.5293 & -0.5774
\end{bmatrix}
\begin{bmatrix}
9.6438 & 0 & 0 \\
0 & 6.3639 & 0 \\
0 & 0 & 0.7056 \\
0 & 0 & 0
\end{bmatrix}
\begin{bmatrix}
-0.7367 & -0.6515 & -0.1811 \\
0.0852 & 0.1762 & -0.9807 \\
0.6708 & -0.7379 & -0.0743
\end{bmatrix}
$$

$$
\boldsymbol{A}_1 = \boldsymbol{u}_1 \boldsymbol{v}_1^\top =
\begin{bmatrix} -0.6710 \\ -0.7197 \\ -0.0939 \\ -0.1515 \end{bmatrix}
\begin{bmatrix} -0.7367 & -0.6515 & -0.1811 \end{bmatrix}
=
\begin{bmatrix}
0.4943 & 0.4372 & 0.1215 \\
0.5302 & 0.4689 & 0.1303 \\
0.0692 & 0.0612 & 0.0170 \\
0.1116 & 0.0987 & 0.0274
\end{bmatrix}
$$

$$
\widehat{\boldsymbol{A}}(2) = \sigma_1 \boldsymbol{A}_1 + \sigma_2 \boldsymbol{A}_2 =
\begin{bmatrix}
4.7801 & 4.2419 & 1.0244 \\
5.2252 & 4.7522 & -0.0250 \\
0.2493 & -0.2743 & 4.9724 \\
0.7495 & 0.2756 & 4.0278
\end{bmatrix}
$$

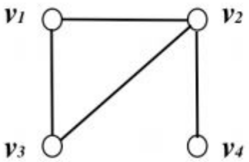
Uygulamalar

- Fotoğraf sıkıştırma
  - Bunun bir örneği yansılarda var.
- Eigenfaces
  - Farklı yüzlerden bir eğitim kümesi oluşturalım
  - SVD ile en dominant özvektörleri bulalım - bu verideki maksimum farkı yaratan doğrultuları verecektir.
  - En büyük *m* tane vektörü seçelim - bu bize yeni bir yüz uzayı tanımlayacak
  - Bütün veriyi bu uzaya eşleyelim
  - Yeni gelen bir yüz için en yakındaki yüzlere bakıp ona göre bir tahmin yapalım

Uygulamalar

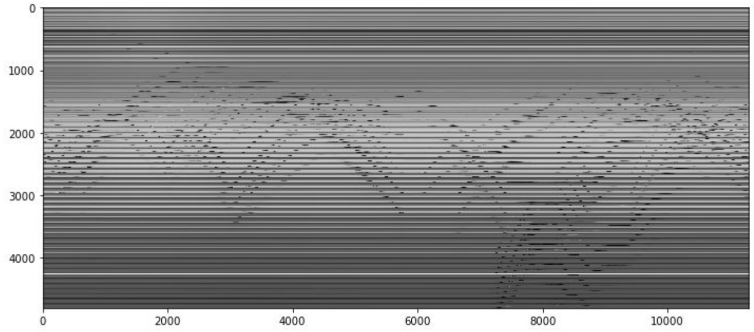
- Kümeleme
  - Uzayda noktalar, ya da noktaların kenarlar ile bağlandığı bir çizge... Yakınlık ya da komşuluk matrisi oluşturulur: **A**
  - Noktaların derece matrisi hesaplanır: **D**
  - **L** = **A** - **D** şeklinde Laplasyen matris oluşturulur ve en büyük **k** özdeğeri hesaplanır
  - Bu özdeğerlere tekabül eden özvektörler ve k-means algoritması ile **k** sınıfa kümelenir.

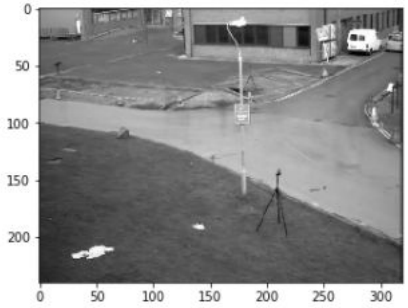
$$\mathbf{L} = \begin{bmatrix} 2 & -1 & -1 & 0 \\ -1 & 3 & -1 & -1 \\ -1 & -1 & 2 & 0 \\ 0 & -1 & 0 & 1 \end{bmatrix}$$



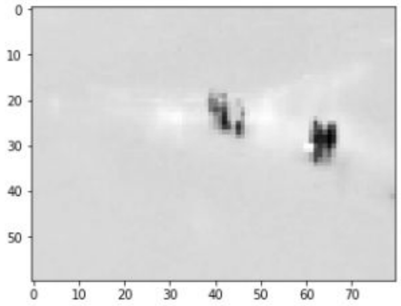
Uygulamalar

- Videolardaki arka planı bulma
    - Bir videoda ön plan, objelerin, insanların, araçların vs hareket ettiği kısımdır
    - Arka plan ise durağandır - üzerinde çok hareket olmaz
    - Videodan frameleri seçip bir bunların sütun olduğu bir matris oluşturalım. Burada görülen dalgalar o piksel üzerinde hareket olduğunu gösterir.
    - Üzerinde hareket olmayan pikseller (satırlar) ise arka planın bir parçasıdır.
    - Dolayısıyla düşük mertebe bir yaklaşım bize arkaplanı verecektir.

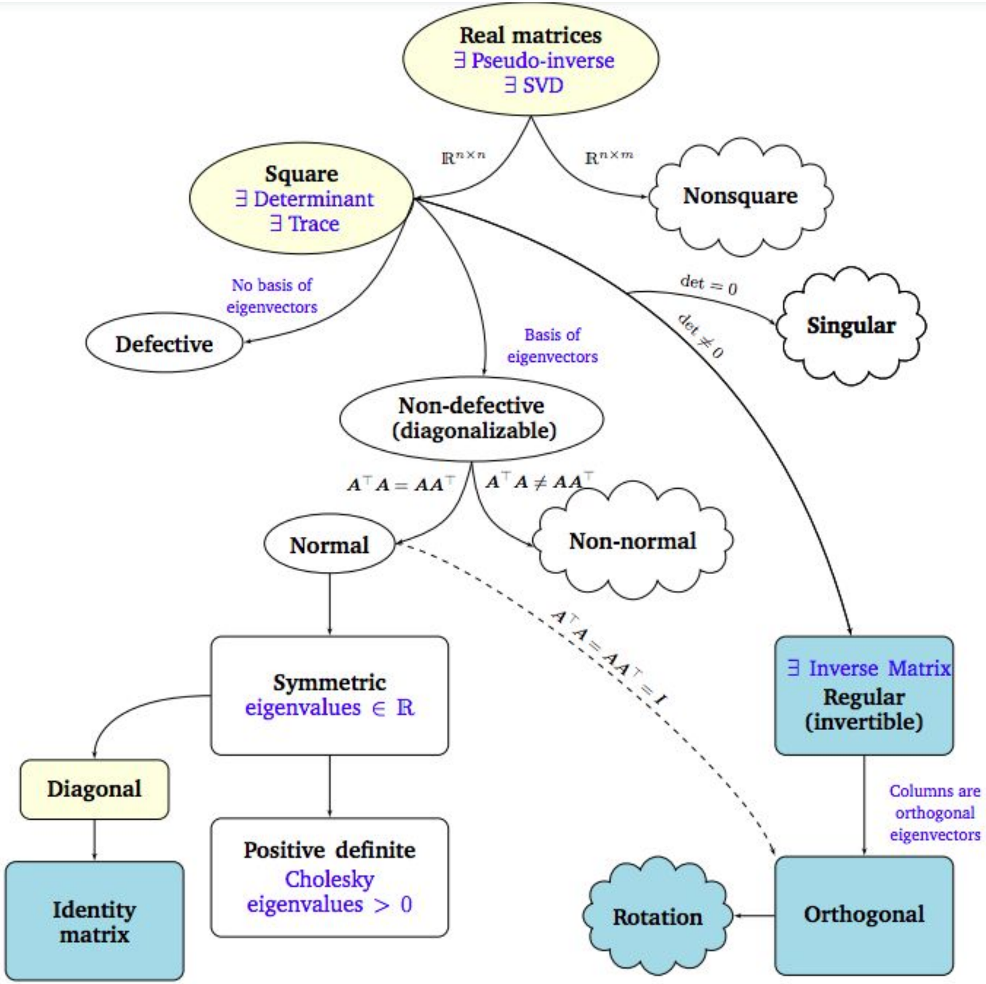




LOW RANK BACKGROUND



FOREGROUND


Real matrices
∃ Pseudo-inverse
∃ SVD
$\mathbb{R}^{n\times n}$
$\mathbb{R}^{n\times m}$
Square
∃ Determinant
∃ Trace
Nonsquare
No basis of eigenvectors
Defective
Basis of eigenvectors
$\det = 0$
$\det \neq 0$
Singular
Non-defective (diagonalizable)
$A^\top A = AA^\top$
$A^\top A \neq AA^\top$
Normal
Non-normal
$A^\top A = AA^\top = I$
Symmetric
eigenvalues $\in \mathbb{R}$
∃ Inverse Matrix
Regular (invertible)
Diagonal
Columns are orthogonal eigenvectors
Identity matrix
Positive definite
Cholesky
eigenvalues > 0
Rotation
Orthogonal

Yararlı kaynaklar

Gilbert Strang: https://www.youtube.com/playlist?list=PLUl4u3cNGP63oMNUHXqIUcrkS2PivhN3k

3Blue1Brown: https://www.youtube.com/playlist?list=PLZHQObOWTQDPD3MizzM2xVFitgF8hE_ab