

MS-A0002 Matriisilaskenta

5. Matriisihajotelmat

Joni Virta, ©Riikka Kangaslampi

Matematiikan ja systeemianalyysin laitos
Aalto-yliopisto

Matriisihajotelmat

Laskentaongelmissa käsiteltävät matriisit ovat tyypillisesti valtavia. Jotta laskut menisivät läpi edes jossain määrin järjellisessä ajassa, käytetään hyväksi mm. matriisihajotelmia.

Ideana niissä on purkaa matriisi useamman, yksinkertaisemman matriisin tuloksi.

Matriisihajotelmia on koko joukko, ne soveltuvat eri käyttötarkoituksiin ja erityyppisille matriiseille.

Tällä kurssilla tarkastelemme varsinaisesti näistä kahta: diagonalisointia ja singulaariarvohajotelmaa (SVD). Lisäksi harjoituksessa 4 esiintyi LU-hajotelma.

5.1 Diagonalisointi

$\mathbb{R}^{n \times n}$ neliömatriisi A , joilla on n riippumatonta ominaisvektoria, voidaan diagonalisoida. Tämä tarkoittaa, että matriisi $A \in \mathbb{R}^{n \times n}$ voidaan kirjoittaa muodossa

$$A = S\Lambda S^{-1},$$

missä matriisin S sarakkeet ovat matriisin A ominaisvektorit, ja matriisi Λ on diagonaalimatriisi, jossa kussakin sarakkeessa on matriisissa S samassa sarakkeessa olevaan ominaisvektoriin liittyvä ominaisarvo.

5.1 Diagonalisointi

Esimerkki 1

Aiemmin laskimme matriisille

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}$$

ominaisarvot ja -vektorit, ja saimme ominaisarvoa -1 vastaavaksi ominaisvektoriksi $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$ ja ominaisarvoa 3 vastaavaksi $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.
Diagonalisoi A .

Vastaus:

$$A = S\Lambda S^{-1} = \begin{pmatrix} 1 & 1 \\ -1 & 1 \end{pmatrix} \begin{pmatrix} -1 & 0 \\ 0 & 3 \end{pmatrix} \begin{pmatrix} 1/2 & -1/2 \\ 1/2 & 1/2 \end{pmatrix}$$

5.1 Diagonalisointi

Huom $n \times n$ -matriisi diagonalisoituu, jos sillä on n lineaarisesti riippumatonta ominaisvektoria. Erityisesti tämä tapahtuu silloin, kun matriisilla on n erisuurta ominaisarvoa, mutta voi siis tapahtua muulloinkin!

Esimerkki 2

Matriisi

$$A = \begin{pmatrix} 1 & -1 \\ 1 & -1 \end{pmatrix}$$

ei diagonalisoidu. Sen ainoa ominaisarvo on nimittäin 0, eli $m_a(0) = 2$. Ominaisvektoreita laskettaessa kuitenkin havaitaan, että löytyy vain yksi ominaisvektori, eli $m_g(0) = 1$, eikä A diagonalisoidu.

5.1 Diagonalisointi

Esimerkki 3

Matriisi

$$A = \begin{pmatrix} 1 & 0 & 2 \\ 0 & 3 & 0 \\ 2 & 0 & 1 \end{pmatrix}$$

diagonalisoituu: Sen ominaisarvot ovat 3 ja -1 , $m_a(3) = 2$ ja $m_a(-1) = 1$, mutta onneksi (kuten aiemmin laskettiin) ominaisarvoa 3 vastaten löytyy ominaistaso, eli saadaan kaksi lin. riippumatonta ominaisvektoria om. arvolle 3. Saadaan hajotelma

$$A = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix} \begin{pmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & -1 \end{pmatrix}^{-1}$$

5.1 Diagonalisointi

Huom. Jos matriisi on diagonalisoituva, sen potensseja on hyvin näppärä laskea:

$$A^2 = S \underbrace{\Lambda S^{-1} S}_{=I} \Lambda S^{-1} = S \Lambda^2 S^{-1}$$

$$A^3 = S \underbrace{\Lambda S^{-1} S}_{=I} \underbrace{\Lambda S^{-1} S}_{=I} \Lambda S^{-1} = S \Lambda^3 S^{-1}$$

$$A^k = S \Lambda^k S^{-1}$$

Diagonaalimatriisin potenssithan ovat helppoja:

$$\Lambda^k = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}^k = \begin{pmatrix} \lambda_1^k & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n^k \end{pmatrix}.$$

5.1 Diagonalisointi

Esimerkki 4

Laske A^{2014} , kun

$$A = \begin{pmatrix} 5 & 1 \\ 0 & 4 \end{pmatrix}$$

Vastaus:

$$A = S \Lambda S^{-1} = \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 5 & 0 \\ 0 & 4 \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix},$$

$$\begin{aligned} A^{2014} &= \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \begin{pmatrix} 5^{2014} & 0 \\ 0 & 4^{2014} \end{pmatrix} \begin{pmatrix} 1 & 1 \\ 0 & -1 \end{pmatrix} \\ &= \begin{pmatrix} 5^{2014} & 5^{2014} - 4^{2014} \\ 0 & 4^{2014} \end{pmatrix} \end{aligned}$$

5.1 Diagonalisointi

Huom. Käänteismatriisin löytäminenkin on diagonalisoidulle matriisille helppoa:

$$A^{-1} = (S\Lambda S^{-1})^{-1} = S\Lambda^{-1}S^{-1}.$$

Diagonaalimatriisillehan

$$\Lambda^{-1} = \begin{pmatrix} \lambda_1 & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n \end{pmatrix}^{-1} = \begin{pmatrix} \lambda_1^{-1} & 0 & 0 \\ 0 & \ddots & 0 \\ 0 & 0 & \lambda_n^{-1} \end{pmatrix}.$$

5.2 Unitaarinen diagonalisointi

Joissakin tapauksissa diagonalisointi onnistuu erityisen mukavasti: tästä on kyse normaalin matriisin unitarisessa diagonalisoinnissa.

Ensin hieman terminologiaa:

- Matriisi $A^* \in \mathbb{C}^{n \times m}$ on matriisin $A \in \mathbb{C}^{m \times n}$ *konjugaattitranspoosi*, jos $(A^*)_{kj} = \overline{A_{jk}}$.
- Matriisi $A \in \mathbb{C}^{n \times n}$ on *symmetrinen*, jos $A^* = A$.
- Matriisi $A \in \mathbb{C}^{n \times n}$ on *normaali*, jos $A^* A = A A^*$.
- Matriisi $U \in \mathbb{C}^{n \times n}$ on *unitaarinen*, jos $U^* = U^{-1}$.

5.2 Unitaarinen diagonalisointi

Lause 5 (Normaalin matriisin unitaaridiagonalisointi)

Matriisille $A \in \mathbb{C}^{n \times n}$ seuraavat ehdot ovat yhtäpitäviä:

- 1** *A on normaali.*
- 2** *$A = U\Lambda U^*$, missä $U \in \mathbb{C}^{n \times n}$ on unitaarinen ja $\Lambda \in \mathbb{C}^{n \times n}$ diagonaalinen.*

Esimerkki 6

Jos A on symmetrinen, eli $A^* = A$, niin A on normaali:

$$A^*A = A^2 = AA^*.$$

Symmetrinen matriisi on siis aina unitaarisesti diagonalisoituva.

5.2 Unitaarinen diagonalisointi

Esimerkki 7

Osoita, että matriisi $A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}$ ei ole normaali.

Lause 8 (Erikoistapaus unitaaridiagonalisoinnista)

Jos $A \in \mathbb{C}^{n \times n}$ on symmetrinen ja sen ominaisarvot ovat erisuuret, niin $A = U\Lambda U^$, missä $U \in \mathbb{C}^{n \times n}$ on unitaarinen ja $\Lambda \in \mathbb{C}^{n \times n}$ diagonaalinen.*

5.2 Unitaarinen diagonalisointi

Huom. Kompleksisille vektoreille $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$ pistetulo on

$$\mathbf{u} \cdot \mathbf{v} = \sum_{i=1}^n u_i \overline{v_i}.$$

Kompleksiset vektorit ovat ortogonaaliset, kun pistetulo on 0.

Huom. Normaalin matriisin eri ominaisarvoja vastaavat normeeratut ominaisvektorit ovat ortonormaaleja. Samaa ominaisarvoa vastaavat ominaisvektorit voidaan valita siten, että ne ovat ortonormaaleita.

Matriisi U saadaan siis normeeratuista, ortonormaaleista ominaisvektoreista.

5.2 Unitaarinen diagonalisointi

Esimerkki 9

Osoita, että matriisi $A = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$ on normaali ja diagonalisoituu unitaarisesti.

Ratkaisu: Lasketaan

$$A^*A = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$$

ja

$$AA^* = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix},$$

joten A on normaali ja se voidaan diagonalisoida unitaarisesti.

5.2 Unitaarinen diagonalisointi

Etsitään ominaisarvot:

$$\det(A - \lambda I) = \begin{vmatrix} -\lambda & 1 \\ -1 & -\lambda \end{vmatrix} = \lambda^2 + 1,$$

joten $\lambda_1 = i$, $\lambda_2 = -i$.

Etsitään ominaisvektorit: yhtälöstä $(A - iI)\mathbf{x} = 0$ saadaan

$$\begin{pmatrix} -i & 1 \\ -1 & -i \end{pmatrix} \sim \begin{pmatrix} -i & 1 \\ 0 & 0 \end{pmatrix}$$

eli ehto $-ix_1 + x_2 = 0$. Valitaan ominaisvektoriksi

$\mathbf{v}_1 = (1/\sqrt{2}, i/\sqrt{2})^T$. (Tällöin sen pituus on 1!)

Vastaavasti ominaisarvolle $\lambda_2 = -i$ saadaan ominaisvektori

$\mathbf{v}_2 = (1/\sqrt{2}, -i/\sqrt{2})^T$.

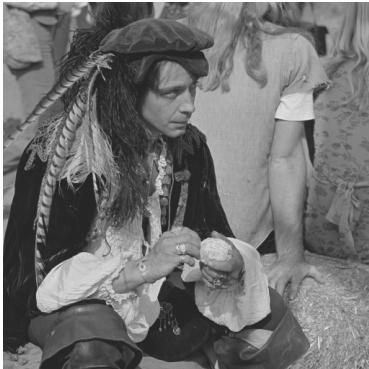
5.2 Unitaarinen diagonalisointi

Matriisille A saadaan siis unitaarinen diagonalisointi

$$\begin{aligned} A &= \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix} \\ &= \begin{pmatrix} 1/\sqrt{2} & 1/\sqrt{2} \\ i/\sqrt{2} & -i/\sqrt{2} \end{pmatrix} \begin{pmatrix} i & 0 \\ 0 & -i \end{pmatrix} \begin{pmatrix} 1/\sqrt{2} & -i/\sqrt{2} \\ 1/\sqrt{2} & i/\sqrt{2} \end{pmatrix} \\ &= U \Lambda U^*. \end{aligned}$$

(Tarkista kertolaskulla!)

5.3 SVD



(a) Alkuperäinen kuva, 463 KB



(b) SVD:llä pakattu kuva, 77 KB

5.3 SVD

Kuten edellä huomattiin, vain neliömatriisin voi diagonalisoida – olettaen, että sillä on täysi määrä riippumattomia ominaisvektoreita. Kaikille matriiseille voidaan kuitenkin tehdä singulaariarvohajotelma (Singular Value Decomposition).

SVD:n ideana on purkaa lineaarikuvaus kolmeen osaan: unitaariseen kuvaukseen, venytykseen ja toiseen unitaariseen kuvaukseen. (Unitaariset kuvaukset voi usein tulkita kierroiksi.)

SVD:tä käytetään mm. signaalinkäsittelyssä ja tilastotieteessä. Siihen perustuvat esimerkiksi eräät kuvanpakkausmenetelmät sekä Googlen hakukoneen toiminta. SVD liittyy läheisesti ns. pääkomponenttianalyysiin, jossa tavoitteena on löytää monidimensioisesta datasta ne komponentit, joilla sen keskeisimmät piirteet voidaan esittää menettämättä oleellista informaatiota.

5.3 SVD

Lause 10 (SVD)

Matriisille $A \in \mathbb{R}^{m \times n}$ on olemassa hajotelma

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

missä U on $m \times m$ -ulotteinen ortogonaalinen matriisi, V on $n \times n$ -ulotteinen ortogonaalinen matriisi ja Σ on $m \times n$ -ulotteinen diagonaalimatriisi, joka sisältää matriisin A singulaariarvot.

Muistutus: Matriisi on ortogonaalinen, jos käänteismatriisi on sen transpoosi: $U^{-1} = U^T$. Tällöin sen sarakevektorit ovat ortonormaaleja. (Reaaliselle matriisille ortogonaalisuus ja unitaarisuus ovat sama asia.)

5.3 SVD

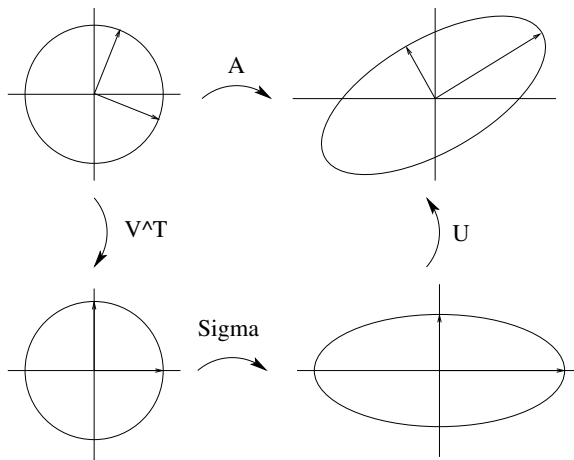


Figure: $A = \text{kierto} + \text{venytys} + \text{kierto}$

5.3 SVD

Huom.

- Vain matriisi Σ on yksikäsitteinen (kun singulaariarvot asetetaan diagonaalille suurimmasta pienimpään), muita voi löytyä useita.
- Luku $\sigma > 0$ on matriisin A singulaariarvo, jos on olemassa vektorit $\mathbf{u}$ ja $\mathbf{v}$ siten, että $A\mathbf{v} = \sigma\mathbf{u}$ ja $A^T\mathbf{u} = \sigma\mathbf{v}$.
- Jos σ on matriisin A singulaariarvo, niin σ^2 on matriisin $A^T A$ ominaisarvo.
- SVD on olemassa myös kompleksisille matriiseille. Tällöin transpoosin V^T sijaan tarvitaan konjugaattitranspoosi V^* ja U ja V ovat unitaarisia, eli $V^{-1} = V^*$, $U^{-1} = U^*$.

5.3 SVD

SVD voidaan laskea seuraavien tietojen perusteella:

- Matriisin Σ diagonaalialkiot σ_j ovat matriisin $A^T A$ ominaisarvojen positiiviset neliöjuuret.
- Matriisin V sarakevektorit $\mathbf{v}_j$ (eli V^T :n rivit) ovat matriisin $A^T A$ yksikköpituiset ominaisvektorit.
- Matriisin U sarakevektorit $\mathbf{u}_j$ ovat matriisin AA^T yksikköpituiset ominaisvektorit.
- Jos $\sigma_j \neq 0$, niin sarakevektori $\mathbf{u}_j = A\mathbf{v}_j/\sigma_j$.

Huom. Matriisit $AA^T \in \mathbb{R}^{m \times m}$ ja $A^T A \in \mathbb{R}^{n \times n}$ ovat aina symmetrisiä, joten niillä on täysi määrä ortogonaalisia ominaisvektoreita. Koko SVD:n idea perustuu tähän.

5.3 SVD

Esimerkki 11

Etsi matriisin

$$A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$$

singulaariarvohajotelma.

Ratkaisu: Matriisin $A^T A = \begin{pmatrix} 10 & 20 \\ 20 & 40 \end{pmatrix}$ ominaisarvot ovat 50 ja 0, joten singulaariarvot ovat $\sigma_1 = \sqrt{50}$, $\sigma_2 = 0$. Matriisin $A^T A$ yksikköpituiset ominaisvektorit ovat $\mathbf{v}_1 = \begin{pmatrix} 1/\sqrt{5} \\ 2/\sqrt{5} \end{pmatrix}$ ja $\mathbf{v}_2 = \begin{pmatrix} 2/\sqrt{5} \\ -1/\sqrt{5} \end{pmatrix}$.

5.3 SVD

Nyt $\mathbf{u}_1 = A\mathbf{v}_1/\sigma_1 = \begin{pmatrix} 5/\sqrt{5} \\ 15/\sqrt{5} \end{pmatrix} \frac{1}{\sqrt{50}} = \begin{pmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{pmatrix}$. Koska $\sigma_2 = 0$, vektoria $\mathbf{u}_2$ ei saada samalla tavalla, vaan se täytyy laskea matriisin AA^T yksikköpituuisena ominaisvektorina: $AA^T = \begin{pmatrix} 5 & 15 \\ 15 & 45 \end{pmatrix}$, ja sen ominaisarvoa 0 vastaava yksikköpituinen ominaisvektori on $\mathbf{u}_2 = \begin{pmatrix} 3/\sqrt{10} \\ -1/\sqrt{10} \end{pmatrix}$. Näin ollen hajotelma on

$$\begin{aligned} A &= U\Sigma V^T \\ &= \begin{pmatrix} 1/\sqrt{10} & 3/\sqrt{10} \\ 3/\sqrt{10} & -1/\sqrt{10} \end{pmatrix} \begin{pmatrix} \sqrt{50} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{pmatrix}. \end{aligned}$$

(Tarkista kertolaskulla!)

5.3 SVD

Alkuperäistä $n \times n$ matriisia voidaan approksimoida SVD:n avulla ottamalla k ensimmäistä saraketta matriisista U , $k \times k$ osa diagonaalimatriisin Σ vasemmasta yläkulmasta ja k ensimmäistä riviä matriisista V^T , eli ensimmäistä k :ta singulaariarvoa vastaava osuus.

Huom: Singulaariarvot oli valittu suuruusjärjestykseen, suurimmasta pienimpään.

Näin saadaan alkuperäistä matriisia approksimoiva matriisi, "jonka rangi on k ".

5.3 SVD

Esimerkki 12

Äskeisen esimerkin matriisille saadaan approksimaatio

$$\begin{aligned} & \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \\ &= \begin{pmatrix} 1/\sqrt{10} & 3/\sqrt{10} \\ 3/\sqrt{10} & -1/\sqrt{10} \end{pmatrix} \begin{pmatrix} \sqrt{50} & 0 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \\ 2/\sqrt{5} & -1/\sqrt{5} \end{pmatrix} \\ &\approx \begin{pmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{pmatrix} \begin{pmatrix} \sqrt{50} \end{pmatrix} \begin{pmatrix} 1/\sqrt{5} & 2/\sqrt{5} \end{pmatrix} \\ &= \begin{pmatrix} 1/\sqrt{10} \\ 3/\sqrt{10} \end{pmatrix} \begin{pmatrix} \sqrt{10} & 2\sqrt{10} \end{pmatrix} \\ &= \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix} \end{aligned}$$

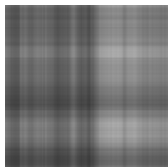
5.3 SVD

Esimerkki 13

Testikuvassa on 512×512 pikseliä, eli se voidaan esittää 512×512 -matriisilla, pikselin väri antaa arvon ko. alkiolle matriisissa. Tälle matriisille voidaan tehdä singulaariarvohajotelma, ja approksimoida sitten kuvaa ottamalla hajotelmasta k :ta ensimmäistä singulaariarvoa vastaava osuus. Huomataan, että vähempikin määrä dataa riittää kuvan esittämiseen tunnistattavasti ja jopa silmämääräisesti riittävän tarkasti.

Seuraavalla sivulla kuvat, kun $k = 1, 10, 20, 50, 100$ ja 200 sekä alkuperäinen kuva.

5.3 SVD

(a) $k = 1$ (b) $k = 10$ (c) $k = 20$ (d) $k = 50$ (e) $k = 100$ (f) $k = 200$ (g) $k = 512$

5.3 SVD

Kun alkuperäisen $m \times n$ -matriisin sijaan käytetään SVD:stä k ensimmäistä singulaariarvoa, tarvittavien lukujen määrä on

$$k + km + kn,$$

eli k singulaariarvoa, niistä vastaavat k ensimmäistä sarakevektoria (pituus m) matriisista U ja k ensimmäistä rivivektoria (pituus n) matriisista V^T .

Neliömatriisille $n \times n$ tämä tarkoittaa, että jos $k < n^2/(1 + 2n)$, niin datan määrä pieneni. Edellä testikuvalle $n = 512$, joten $k < 255$ vähentää datamäärää.

5.3 SVD

Esimerkki 14

On saatu seuraava singulaariarvohajotelma:

$$\begin{pmatrix} 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 \\ 1 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 4 & 0 & 0 & 0 & 0 \\ 0 & 3 & 0 & 0 & 0 \\ 0 & 0 & \sqrt{5} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 1/\sqrt{5} & 0 & 0 & 0 & 2/\sqrt{5} \\ 0 & 0 & 0 & 1 & 0 \\ -2/\sqrt{5} & 0 & 0 & 0 & 1/\sqrt{5} \end{pmatrix}$$

Approksimoi kyseessä olevaa matriisia kahden ja sitten kolmen singulaariarvon avulla. Mitä saat? Miksi? Laske myös alkuperäinen matriisi.

5.3 SVD

Esimerkki 15

Käy läpi SVD:hen perustuva hahmontunnistusdemo Matlabilla. Demon ohjeet ovat jaossa kurssin MyCourses-sivulla.

Huomattiin siis, että SVD:n avulla voidaan poimia datasta olennainen ja päästä turhasta “kohinasta” eroon, pienentäen samalla käsiteltävää datamäärää merkittävästi.

SVD:n muut sovellukset jätämme myöhempisiin kursseihin, tämä kurssi päättyy tähän!

Nämä kalvot perustuvat Riikka Kangaslammen alkuperäisiin, joita ovat myöhemmin muokanneet Mikael Laaksonen ja Joni Virta.