...

Robert Susmaga
Instytut Informatyki
ul. Piotrowo 2
Poznań

kontakt mail'owy
Robert.Susmaga@CS.PUT.Poznan.PL
kontakt osobisty
Centrum Wykładowe, „blok informatyki”, pok. 7

Wyłączenie odpowiedzialności

Prezentowane materiały, będące dodatkiem pomocniczym do wykładów, z konieczności fragmentarycznym i niedopracowanym, należy wykorzystywać z pełną świadomością faktu, że mogą nie być pozbawione przypadkowych błędów, braków, wypaczeń i przeinaczeń :-)

Autor

…

Rozkład EVD

- Widmo spec(**A**) macierzy $\mathbf{A}_{nxn}$ może zawierać mniej elementów niż n (ponieważ jest zbiorem, a zbiory nie zawierają powtórzeń)
  – w celu ujednoznacznienia wyniku definiuje się tzw. krotności elementów
    - są to oczywiście krotności rozwiązań wielomianu charakterystycznego
  – zapis alternatywny:
    zastosowanie wektora (a nie zbioru) wartości własnych

Rozkład EVD

- Zbiór wartości własnych spec(**A**)
  a wektor wartości własnych eig(**A**)
  i posortowany wektor wartości własnych eigs(**A**)
  - sortowanie względem wartości bezwzględnych
  - przykłady dla macierzy $\mathbf{A}_{4x4}$
    - spec(**A**) = {1, –2, 3, 4}, eig(**A**) = [1, –2, 3, 4], eigs(**A**) = [4, 3, –2, 1],
    - spec(**A**) = {1, 2, 3}, eig(**A**) = [3, 2, 3, 1], eigs(**A**) = [3, 3, 2, 1],
    - spec(**A**) = {1, –3}, eig(**A**) = [1, 1, –3, 1], eigs(**A**) = [–3, 1, 1, 1],
    - spec(**A**) = {3}, eig(**A**) = [3, 3, 3, 3], eigs(**A**) = [3, 3, 3, 3],

Rozkład EVD

- W praktyce to wektor (a nie zbiór!) wartości własnych jest używany w rozkładzie EVD
  - jest tak, ponieważ macierzowy układ równań **AK** = **KL** wymaga macierzy **K** i **L** o rozmiarach nxn
    - potrzebujemy więc
      - n wartości własnych
      - n wektorów własnych
  - kolejność kolumn macierzy **K** i elementów diagonalnych macierzy **L** musi być „zgodna”
    - wektory własne muszą odpowiadać wartościom własnym

Rozkład EVD

- Dodatkowo, podczas gdy wartości własne są ustalone „bezwzględnie”, wektory własne są ustalone „względnie”
  - jeżeli **k** jest wektorem własnym macierzy **A**, to d**k**, gdzie d jest niezerowym skalarem, jest także wektorem własnym tej macierzy (o innej długości)
    - d**k** odpowiada oczywiście tej samej wartości własnej, co **k**
  - dzięki temu generowane w praktyce wektory własne mogą być odpowiednio dobierane (np. w taki sposób, aby miały zadaną długość)

Rozkład EVD

- Idea rozkładu EVD
  - niech **AK** = **KL**, gdzie **L** jest macierzą diagonalną
  - czyli
    - elementy przekątnej macierzy **L** stanowią wartości własne macierzy **A**
    - kolumny macierzy **K** stanowią wektory własne macierzy **A**
  - jeżeli dodatkowo det(**K**) $\neq$ 0, to wtedy

    $\mathbf{AK} = \mathbf{KL}$

    $\mathbf{AKK}^{-1} = \mathbf{KLK}^{-1}$

    $\mathbf{A} = \mathbf{KLK}^{-1}$
  - czyli
    - $\mathbf{KLK}^{-1}$ jest rozkładem EVD macierzy **A**

Rozkład EVD

- Jednoznaczność rozkładu EVD

  AK = **KL**

  – niech **D** będzie macierzą diagonalną
    - wobec jednoczesnej diagonalności macierzy **L** mamy: **LD** = **DL**

  – wtedy

  AKD = **KLD**

  AKD = **KDL**

  A(**KD**) = (**KD**)**L**

  – czyli
    - elementy przekątnej macierzy **L** stanowią wartości własne macierzy **A**
    - kolumny macierzy **KD** stanowią wektory własne macierzy **A**

Rozkład EVD

- Jednoznaczność rozkładu EVD, c.d.

  $\mathbf{AKD} = \mathbf{KDL}$

  – jeżeli $\det(\mathbf{D}) \neq 0$, to $\det(\mathbf{KD}) = \det(\mathbf{K})\det(\mathbf{D}) \neq 0$, i wtedy

  $\mathbf{A}(\mathbf{KD})(\mathbf{KD})^{-1} = (\mathbf{KD})\mathbf{L}(\mathbf{KD})^{-1}$

  $\mathbf{A} = (\mathbf{KD})\mathbf{L}(\mathbf{KD})^{-1}$

Rozkład EVD

- Jednoznaczność rozkładu EVD, c.d.
  - pytanie: a czy uogólnienie jest możliwe, gdy $\mathbf{D}$ jest nieosobliwa ale niekoniecznie diagonalna (oznaczenie $\mathbf{N}$)?
    - czyli: czy dla nieosobliwych $\mathbf{N}$ zachodzi $\mathbf{A} = (\mathbf{KN})\mathbf{L}(\mathbf{KN})^{-1}$?
  - odpowiedź: tak, o ile $\mathbf{L}$ i $\mathbf{N}$ komutują, tzn. $\mathbf{LN} = \mathbf{NL}$
  - uzasadnienie

    $\mathbf{AK} = \mathbf{KL}$

    $\mathbf{AKN} = \mathbf{KLN}$

    $\mathbf{AKN} = \mathbf{KNL}$

    $\mathbf{AKN}(\mathbf{KN})^{-1} = (\mathbf{KN})\mathbf{L}(\mathbf{KN})^{-1}$

    $\mathbf{A} = (\mathbf{KN})\mathbf{L}(\mathbf{KN})^{-1}$

Rozkład EVD

- Jednoznaczność rozkładu EVD, c.d.
  - pytanie: a czy uogólnienie jest możliwe, gdy nieosobliwa macierz zostanie wykorzystana w mnożeniu lewostronnym (oznaczenie **M**)?
    - czyli: czy dla nieosobliwych **M** zachodzi $\mathbf{A} = (\mathbf{MK})\mathbf{L}(\mathbf{MK})^{-1}$?
  - odpowiedź: tak, o ile **M** i **A** komutują, tzn. $\mathbf{MA} = \mathbf{AM}$
  - uzasadnienie

$\mathbf{AK} = \mathbf{KL}$

$\mathbf{MAK} = \mathbf{MKL}$

$\mathbf{AMK} = \mathbf{MKL}$

$\mathbf{AMK}(\mathbf{MK})^{-1} = (\mathbf{MK})\mathbf{L}(\mathbf{MK})^{-1}$

$\mathbf{A} = (\mathbf{MK})\mathbf{L}(\mathbf{MK})^{-1}$

Rozkład EVD

- Szczególne źródło niejednoznaczności:
  - kolejność wartości własnych (na przekątnej macierzy diagonalnej **L**)

  oraz
  - odpowiadających im wektorów własnych (w macierzy **K**)

Rozkład EVD

- Macierz $\mathbf{P}_{nxn} = [p_{ij}]$ nazywamy macierzą permutacyjną, gdy
  - $\forall_i \forall_j \, p_{ij} \in \{0, 1\}$
    - elementami są wyłącznie 0 i 1
  - $\forall_i \sum_j p_{ij} = 1$
    - suma każdego wiersza równa 1
  - $\forall_j \sum_i p_{ij} = 1$
    - suma każdej kolumny równa 1
  - uwagi
  - dokładnie jedna jedynka w każdej kolumnie i w każdym wierszu
  - macierz permutacyjna **P** jest macierzą ortogonalną
    - patrz: dalsze slajdy

Rozkład EVD

- Działanie macierzy permutacyjnej: jeżeli **P** jest macierzą permutacyjną, to
  - operacja **XP** ustawia kolumny macierzy **X**

  w takiej samej kolejności, w jakiej
  - operacja $\mathbf{P}^T\mathbf{X}$ ustawia wiersze macierzy **X**

  i w takiej samej kolejności, w jakiej
  - operacja $\mathbf{P}^T\mathbf{XP}$ ustawia wiersze i kolumny macierzy **X**
    - w szczególności
      - operacja $\mathbf{P}^T\mathbf{LP}$ ustawia elementy diagonalne diagonalnej macierzy **L** (a więc $\mathbf{P}^T\mathbf{LP}$ jest także macierzą diagonalną)

Rozkład EVD

- Niech **AK** = **KL**
  - przy det(**K**) $\neq$ 0 zachodzi także **A** = **KLK**$^{-1}$

  i niech **P** będzie macierzą permutacyjną
  - jeżeli kolumny macierzy **K** i elementy diagonalne macierzy **L** są „zgodne”, to kolumny macierzy **KP** i elementy diagonalne macierzy **P**^{T}**LP** są także „zgodne”

Rozkład EVD

- Wtedy

  $\mathbf{AK} = \mathbf{KL}$

  $\mathbf{AKP} = \mathbf{KLP}$

  $\mathbf{AKP} = \mathbf{KILP}$

  $\mathbf{AKP} = \mathbf{KPP}^{\mathrm{T}}\mathbf{LP}$

  $\mathbf{AKP} = \mathbf{KP}(\mathbf{P}^{\mathrm{T}}\mathbf{LP})$

  – jeżeli $\det(\mathbf{K}) \neq 0$, to $\det(\mathbf{KP}) = \det(\mathbf{K})\det(\mathbf{P}) \neq 0$, i wtedy

  $\mathbf{AKP}(\mathbf{KP})^{-1} = \mathbf{KP}(\mathbf{P}^{\mathrm{T}}\mathbf{LP})(\mathbf{KP})^{-1}$

  $\mathbf{A} = \mathbf{KP}(\mathbf{P}^{\mathrm{T}}\mathbf{LP})(\mathbf{KP})^{-1}$

Rozkład EVD

- Jednoznaczność rozkładu EVD
  - powyższe uogólnienia można składać, np. dysponując
    - macierzą permutacyjną **P**,
    - nieosobliwą macierzą **N** komutującą z macierzą $\mathbf{P}^T\mathbf{L}\mathbf{P}$
    - nieosobliwą macierzą **M** komutującą z macierzą **A**

  spełnione są zależności

  $\mathbf{AK} = \mathbf{KL}$

  ...

  $\mathbf{A}(\mathbf{KP}) = (\mathbf{KP})(\mathbf{P}^T\mathbf{LP})$

  ...

  $\mathbf{A}(\mathbf{KPN}) = (\mathbf{KPN})(\mathbf{P}^T\mathbf{LP})$

  ...

  $\mathbf{A}(\mathbf{MKPN}) = (\mathbf{MKPN})(\mathbf{P}^T\mathbf{LP})$

  które, przy odwracalności **K**, **P**, **N** i **M**,
  służą do tworzenia ogólniejszych rozkładów EVD

…

Interpretacja geometryczna iloczynu skalarnego

- Elementy przestrzeni liniowej n-wymiarowej (wektory)
  - interpretacja wektora zerowego $\mathbf{0} = [0, \ldots, 0]^T$
    - punkt: środek układu współrzędnych
  - interpretacje wektorów niezerowych
    - punkty: $\mathbf{x} = [x_1, \ldots, x_n]^T$
    - strzałki: $\mathbf{x} \equiv \mathbf{0} \rightarrow \mathbf{x}$ (początek w punkcie **0**, koniec w punkcie **x**)
- Rozszerzone interpretacje
  - wektory swobodne
    - elementy przestrzeni afinicznej

Interpretacja geometryczna iloczynu skalarnego

- Podstawowe operacje wektorowe
  - mnożenie przez skalar
  - dodawanie
  - odejmowanie
- Definicje
- Interpretacje geometryczne

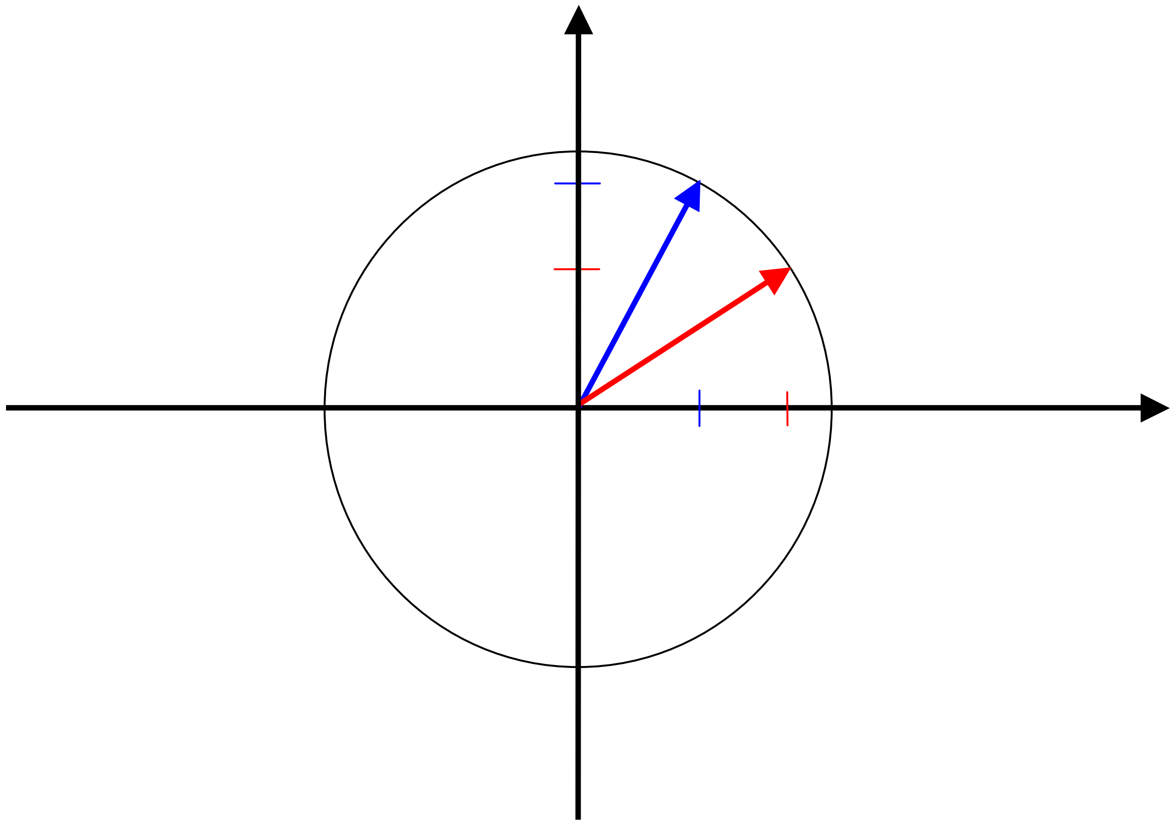
Interpretacja geometryczna iloczynu skalarnego

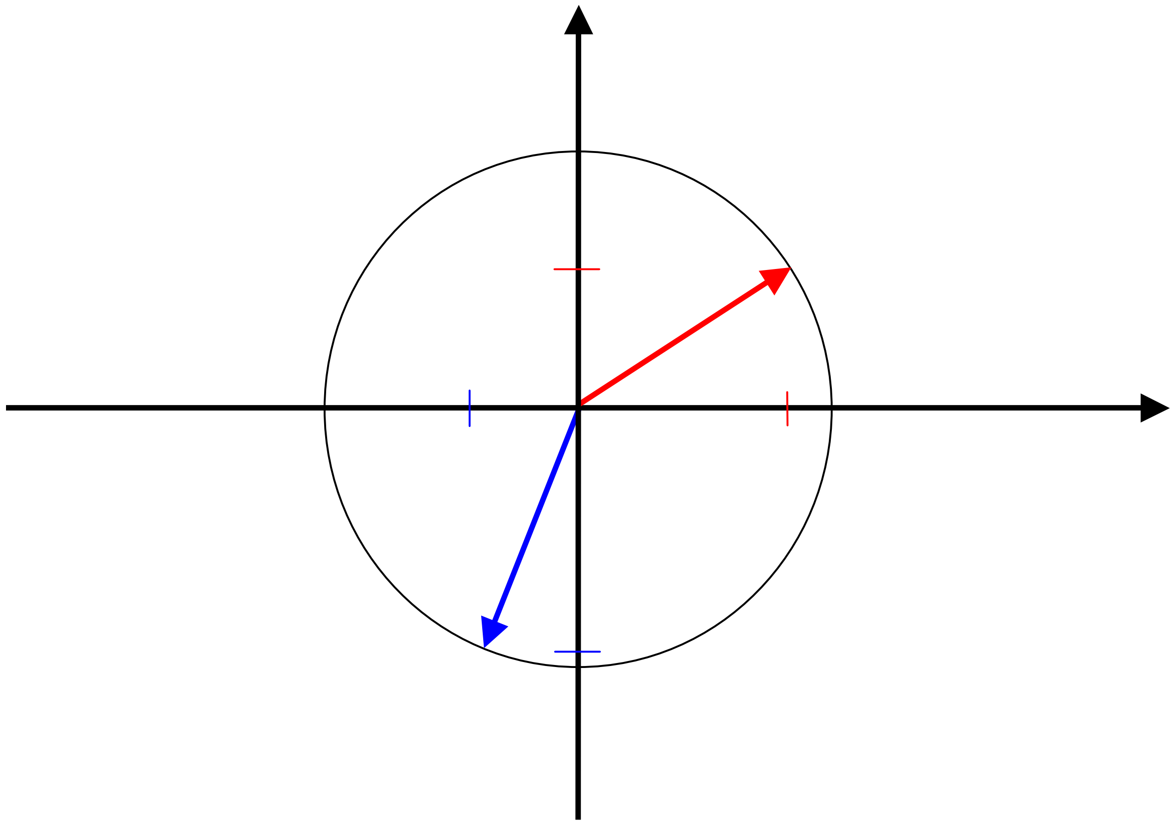
- Iloczyn skalarny wektorów
  - iloczynem skalarnym wektorów $\mathbf{x} = [x_1, \ldots, x_n]^T$ oraz $\mathbf{y} = [y_1, \ldots, y_n]^T$ nazywamy wartość $\Sigma_i x_i y_i$
  - iloczyn skalarny w zapisie
    - macierzowym: $\mathbf{x}^T\mathbf{y}$ (jako mnożenie macierzy)
    - fizycznym: $\langle \mathbf{x},\mathbf{y} \rangle$ (stosowanym w fizyce)
  - oczywiście zachodzi: $\mathbf{x}^T\mathbf{y} = [\Sigma_i x_i y_i] = \Sigma_i x_i y_i$
  - podstawowe właściwości
    - wynik operacji $\mathbf{x}^T\mathbf{y}$ jest macierzą o rozmiarach 1x1, a więc jest macierzą symetryczną
      - w rezultacie powyższej symetrii zachodzi $\mathbf{x}^T\mathbf{y} = (\mathbf{x}^T\mathbf{y})^T = \mathbf{y}^T(\mathbf{x}^T)^T = \mathbf{y}^T\mathbf{x}$
    - wynik operacji $\mathbf{x}^T\mathbf{y}$ jest macierzą o rozmiarach 1x1, a więc może być utożsamiany ze skalarem $\Sigma_i x_i y_i$
      - w rezultacie można zapisać $\mathbf{x}^T\mathbf{y}\cdot\mathbf{A} = \mathbf{A}\cdot\mathbf{x}^T\mathbf{y}$ (przemienność mnożenia przez skalar)

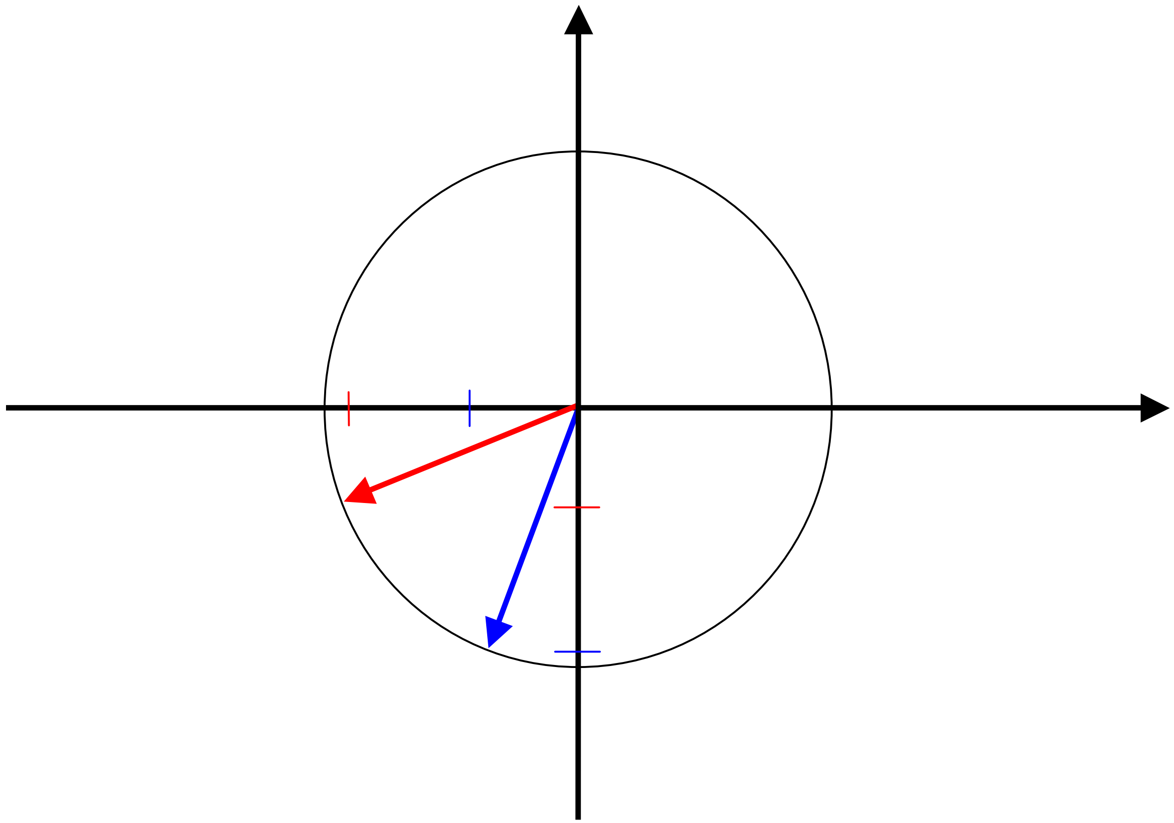
Interpretacja geometryczna iloczynu skalarnego

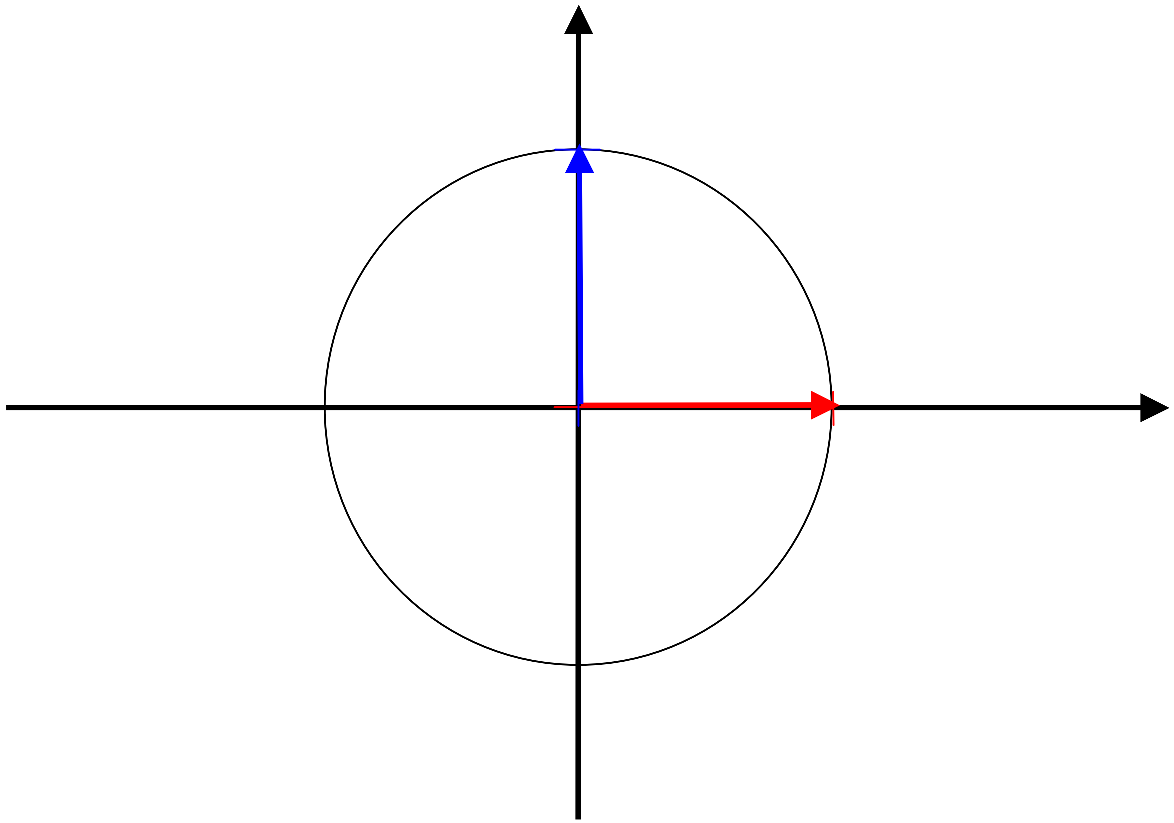
- Właściwości iloczynu skalarnego wektorów **x** i **y**
  - właściwości iloczynu dla dowolnych wektorów
    - $\mathbf{x} = \mathbf{0} \vee \mathbf{y} = \mathbf{0} \Rightarrow \mathbf{x}^T\mathbf{y} = 0$

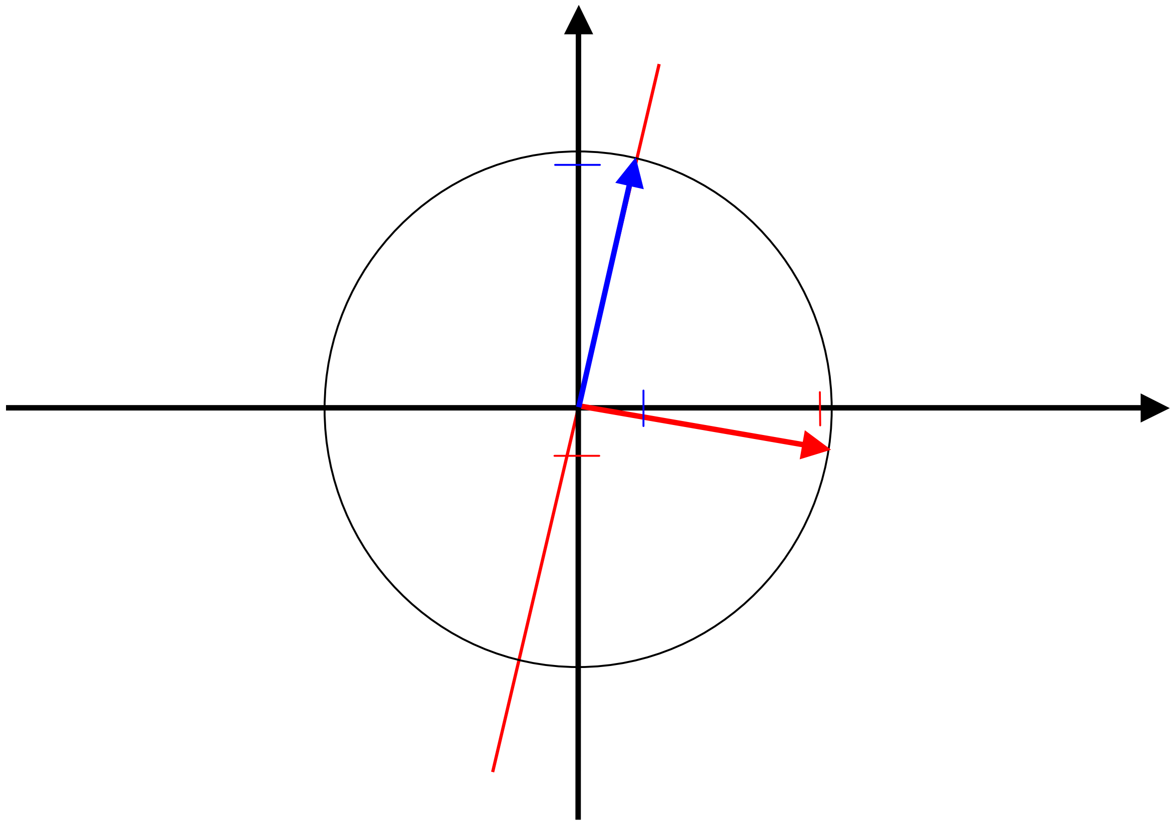
    (implikacja w drugą stronę nie zachodzi)
  - właściwości iloczynu dla niezerowych wektorów
    - czy możliwa jest sytuacja $\mathbf{x}^T\mathbf{y} = 0$, a jeżeli tak, to kiedy?











Interpretacja geometryczna iloczynu skalarnego

- Kąt pomiędzy niezerowymi wektorami **x** i **y** jest
  - ostry wtedy i tylko wtedy, gdy $\mathbf{x}^T\mathbf{y} > 0$
  - prosty wtedy i tylko wtedy, gdy $\mathbf{x}^T\mathbf{y} = 0$ (wektory ortogonalne)
  - rozwarty wtedy i tylko wtedy, gdy $\mathbf{x}^T\mathbf{y} < 0$
- Kąt pomiędzy wektorami **x** i **y**, z których przynajmniej jeden jest zerowy, nie jest określony
  - nadal jednak stosuje się pojęcie ortogonalności, tzn.: jeżeli $\mathbf{x}^T\mathbf{y} = 0$, to **x** i **y** nazywamy ortogonalnymi
  - wnioski:
    - każdy wektor jest ortogonalny do wektora **0**
    - wektor **0** jest ortogonalny do każdego wektora

…

Interpretacja geometryczna normy

- Zadaniem normy wektorowej, oznaczenie $||\mathbf{x}||$, gdzie **x** jest dowolnym wektorem, jest mierzenie „skali wielkości” wektora **x**
  – norma jest więc wartością nieujemną:
    - czyli: $||\mathbf{x}|| \geq 0$

    która mierzy pewną „odległość” wektora od wektora zerowego
    - stąd: podstawowa cecha normy wektorowej: $||\mathbf{0}|| = 0$

  – norma powinna także spełniać aksjomaty zależności liniowej
- Ostatecznie zakładane właściwości normy wektorowej:

  $||\mathbf{x}|| \geq 0$, przy czym $||\mathbf{x}|| = 0 \;\Leftrightarrow\; \mathbf{x} = \mathbf{0}$

  $||a\mathbf{x}|| = |a|\cdot||\mathbf{x}||$

  $||\mathbf{x}+\mathbf{y}|| \leq ||\mathbf{x}|| + ||\mathbf{y}||$ (nierówność trójkąta)

Interpretacja geometryczna normy

- Podstawowa norma wektorowa (Euklidesa)

  $$||\mathbf{x}||_2 = (\Sigma_i|x_i|^2)^{1/2} = (\mathbf{x}^T\mathbf{x})^{1/2}$$

  – norma spełnia założone właściwości norm wektorowych:

  $||\mathbf{x}||_2 \geq 0$, przy czym $||\mathbf{x}||_2 = 0 \;\Leftrightarrow\; \mathbf{x} = \mathbf{0}$

  $||a\mathbf{x}||_2 = |a|\cdot||\mathbf{x}||_2$

  $||\mathbf{x}+\mathbf{y}||_2 \leq ||\mathbf{x}||_2 + ||\mathbf{y}||_2$
- Interpretacja normy Euklidesa:
  długość reprezentującej wektor strzałki

Interpretacja geometryczna normy

- Uogólnienie normy Euklidesa, definiowana dla $p \geq 1$, p-norma:

  $$||\mathbf{x}||_p = (\Sigma_i |x_i|^p)^{1/p}$$

  – p-norma spełnia założone właściwości norm wektorowych:

  $||\mathbf{x}||_p \geq 0$, przy czym $||\mathbf{x}||_p = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$

  $||a\mathbf{x}||_p = |a| \cdot ||\mathbf{x}||_p$

  $||\mathbf{x}+\mathbf{y}||_p \leq ||\mathbf{x}||_p + ||\mathbf{y}||_p$

- Najczęściej stosowane wersje p-normy:

  $||\mathbf{x}||_1 = \Sigma_i |x_i|$ – 1-norma, norma manhattańska

  $||\mathbf{x}||_2 = (\Sigma_i |x_i|^2)^{1/2}$ – 2-norma, norma Euklidesa

  $||\mathbf{x}||_\infty = \lim_{p \to \infty} (\Sigma_i |x_i|^p)^{1/p} = \max_i |x_i|$ – ∞-norma, norma Czebyszewa

Interpretacja geometryczna normy

- Uwaga: norma Euklidesa będzie stosowana wyłącznie w dalszej części wykładu
  - a więc: $||\mathbf{x}||$ należy zawsze rozumieć jako $||\mathbf{x}||_2$

Interpretacja geometryczna normy

- Wektory unormowane
  - wektor **x** nazywamy unormowanym, gdy ||**x**|| = 1
  - każdy niezerowy wektor **x** można przekształcić w unormowany wektor **y** = **x**/||**x**||
  - każdy wektor postaci **x**/||**x**|| jest wektorem unormowanym:
    - wykorzystując ||a**x**|| = |a|·||**x**|| mamy
      || **x**/||**x**|| || = || (1/||**x**||)·**x** || = |(1/||**x**||)|·||**x**||
    - ale zachodzi również ||**x**|| ≥ 0, w rezultacie czego |(1/||**x**||)| = 1/||**x**||
    - ostatecznie
      || **x**/||**x**|| || = |(1/||**x**||)|·||**x**|| = (1/||**x**||)·||**x**|| = ||**x**||/||**x**|| = 1

…

Ortogonalność i ortonormalność

- Wektor **x** jest unormowany, gdy
  - $\mathbf{x}^T\mathbf{x} = 1$
- Dwa wektory (**x** i **y**) są ortogonalne, gdy
  - $\mathbf{x}^T\mathbf{y} = 0$
- Zbiór wektorów $\{\mathbf{x}_1, ..., \mathbf{x}_n\}$ nazywamy zbiorem wektorów ortogonalnych, gdy
  - dla każdej pary indeksów i oraz j zachodzi $i \neq j \;\Rightarrow\; (\mathbf{x}_i)^T\mathbf{x}_j = 0$ (ortogonalność parami)
  - zbiór wektorów ortogonalnych jest zbiorem wektorów niezależnych
- Zbiór wektorów $\{\mathbf{x}_1, ..., \mathbf{x}_n\}$ nazywamy zbiorem wektorów ortonormalnych, gdy
  - dla każdej pary indeksów i oraz j zachodzi
    - $i = j \;\Rightarrow\; (\mathbf{x}_i)^T\mathbf{x}_j = 1$ (unormowanie)
    - $i \neq j \;\Rightarrow\; (\mathbf{x}_i)^T\mathbf{x}_j = 0$ (ortogonalność)

    (unormowane wektory, które są wzajemnie ortogonalne)
  - zbiór wektorów ortonormalnych jest zbiorem wektorów niezależnych

Ortogonalność i ortonormalność

- Macierz jest macierzą ortogonalną, gdy jej kolumny stanowią zbiór wektorów ortonormalnych
  - niestety, nie stosuje się zbyt powszechnie (bardziej konsekwentnej) nazwy „macierz ortonormalna”

Ortogonalność i ortonormalność

- Podstawowa właściwość macierzy ortogonalnej
  - niech $\mathbf{K}_{nxn} = [\mathbf{k}_1, \mathbf{k}_2, ..., \mathbf{k}_n]$ będzie macierzą (kwadratową), której kolumny stanowią zbiór wektorów ortonormalnych
  - wtedy

$$\mathbf{K}^T\mathbf{K} = \begin{bmatrix} (\mathbf{k}_1)^T\mathbf{k}_1 & (\mathbf{k}_1)^T\mathbf{k}_2 & ... & (\mathbf{k}_1)^T\mathbf{k}_n \\ (\mathbf{k}_2)^T\mathbf{k}_1 & (\mathbf{k}_2)^T\mathbf{k}_2 & ... & (\mathbf{k}_2)^T\mathbf{k}_n \\ ... & ... & ... & ... \\ (\mathbf{k}_n)^T\mathbf{k}_1 & (\mathbf{k}_n)^T\mathbf{k}_2 & ... & (\mathbf{k}_n)^T\mathbf{k}_n \end{bmatrix} = \begin{bmatrix} 1 & 0 & ... & 0 \\ 0 & 1 & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \end{bmatrix} = \mathbf{I}$$

Ortogonalność i ortonormalność

- Odwrotność macierzy kwadratowej, której kolumny stanowią zbiór wektorów ortonormalnych
  - ponieważ $\mathbf{K}^T\mathbf{K} = \mathbf{I}$, więc
    - macierz $\mathbf{K}^T$ jest odwrotnością macierzy $\mathbf{K}$, a więc $\mathbf{K}^{-1} = \mathbf{K}^T$
    - macierz $\mathbf{K}$ jest odwrotnością macierzy $\mathbf{K}^T$, a więc $(\mathbf{K}^T)^{-1} = \mathbf{K}$
- Zaskakująca (i niezwykle przydatna!) właściwość
  - odwrotnością macierzy ortogonalnej jest jej transpozycja

Macierze ortogonalne

- Macierz **U** nazywamy ortogonalną, gdy $\mathbf{U}^T\mathbf{U} = \mathbf{U}\mathbf{U}^T = \mathbf{I}$
  - uwagi

Macierze ortogonalne

- Jeżeli macierze **U** i **V** są ortogonalne, to macierz **UV** jest ortogonalna
  - uzasadnienie
  - niech **U** i **V** będą macierzami ortogonalnymi i niech **W** = **UV**
  - *macierz* $\boldsymbol{K}$ *nazywamy ortogonalną, gdy* $\boldsymbol{K}^T\boldsymbol{K} = \boldsymbol{K}\boldsymbol{K}^T = \boldsymbol{I}$ *[ref]*
  - a więc macierze **U** i **V** spełniają: $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ oraz $\mathbf{V}^T\mathbf{V} = \mathbf{I}$
  - jednocześnie
    $\mathbf{W}^T\mathbf{W} = (\mathbf{UV})^T\mathbf{UV} = \mathbf{V}^T\mathbf{U}^T\mathbf{UV} = \mathbf{V}^T\mathbf{IV} = \mathbf{V}^T\mathbf{V} = \mathbf{I}$
  - wniosek: macierz **UV** jest ortogonalna ■

Macierze ortogonalne

- Jeżeli $\mathbf{U}$ jest ortogonalna, to $\det(\mathbf{U}) = \pm 1$
  - uzasadnienie
  - macierz ortogonalna $\mathbf{U}$ spełnia zależność $\mathbf{U}^T\mathbf{U} = \mathbf{I}$, a więc

    $\det(\mathbf{U}^T\mathbf{U}) = \det(\mathbf{I}) = 1$
  - jednocześnie ponieważ dla każdej macierzy kwadratowej zachodzi $\det(\mathbf{A}^T) = \det(\mathbf{A})$ więc

    $\det(\mathbf{U}^T\mathbf{U}) = \det(\mathbf{U}^T)\det(\mathbf{U}) = \det(\mathbf{U})\det(\mathbf{U}) = \det(\mathbf{U})^2$
  - czyli

    $\det(\mathbf{U})^2 = 1$
  - wniosek: $\det(\mathbf{U}) = \pm 1$

Macierze ortogonalne

- Jeżeli **U** jest macierzą ortogonalną i λ jest jej wartością własną, to $|\lambda| = 1$
  - uzasadnienie
  - niech λ będzie wartością własną a **k** odpowiadającym jej wektorem własnym macierzy **U**, tzn. $\mathbf{Uk} = \lambda\mathbf{k}$ i $\mathbf{k} \neq \mathbf{0}$
  - wtedy $\mathbf{k} \neq \mathbf{0} \;\Leftrightarrow\; \mathbf{k}^T\mathbf{k} \neq 0$ i jednocześnie

$$
\begin{aligned}
\mathbf{Uk} = \lambda\mathbf{k} \;&\Rightarrow\; (\mathbf{Uk})^T\mathbf{Uk} = (\lambda\mathbf{k})^T(\lambda\mathbf{k}) \;\Leftrightarrow \\
&\Leftrightarrow\; \mathbf{k}^T\mathbf{U}^T\mathbf{Uk} = \mathbf{k}^T\lambda^T\lambda\mathbf{k} \;\Leftrightarrow \\
&\Leftrightarrow\; \mathbf{k}^T\mathbf{Ik} = \mathbf{k}^T\lambda\lambda\mathbf{k} \;\Leftrightarrow \\
&\Leftrightarrow\; \mathbf{k}^T\mathbf{k} = \mathbf{k}^T\lambda^2\mathbf{k} \;\Leftrightarrow \\
&\Leftrightarrow\; \mathbf{k}^T\mathbf{k} = \lambda^2\mathbf{k}^T\mathbf{k} \;\Rightarrow \\
&\Rightarrow\; \mathbf{k}^T\mathbf{k}/\mathbf{k}^T\mathbf{k} = \lambda^2 \;\Leftrightarrow \\
&\Rightarrow\; 1 = \lambda^2 \;\Leftrightarrow \\
&\Leftrightarrow\; |\lambda| = 1
\end{aligned}
$$

  - wniosek:
    moduły wartości własnych macierzy ortogonalnej są równe jeden

Macierze ortogonalne

- Jeżeli $\mathbf{U}^T\mathbf{U} = \mathbf{I}$ i $\mathbf{y} = \mathbf{U}\mathbf{x}$, to $||\mathbf{y}|| = ||\mathbf{x}||$
  - uzasadnienie
    $||\mathbf{y}|| = ||\mathbf{U}\mathbf{x}|| = ((\mathbf{U}\mathbf{x})^T\mathbf{U}\mathbf{x})^{1/2} = (\mathbf{x}^T\mathbf{U}^T\mathbf{U}\mathbf{x})^{1/2} = (\mathbf{x}^T\mathbf{I}\mathbf{x})^{1/2} = (\mathbf{x}^T\mathbf{x})^{1/2} = ||\mathbf{x}||$
  - wniosek: $||\mathbf{y}|| = ||\mathbf{x}||$

…

Miara kąta pomiędzy wektorami

- Właściwości cosinusa kąta α z przedziału $[0, \pi]$
  - $\cos(\alpha) \in [-1,+1]$, przy czym
    - $\cos(\alpha) = +1$ dla $\alpha = 0$
    - $\cos(\alpha) = 0$ dla $\alpha = \pi/2$
    - $\cos(\alpha) = -1$ dla $\alpha = \pi$

Miara kąta pomiędzy wektorami

- Właściwości funkcji cosinus w ustalonym przedziale
  - ponieważ
    - przy $x \in [0, \pi]$ zachodzi $\cos(x) \in [-1,+1]$

    oraz
    - funkcja $y = \cos(x)$ jest odwracalna
      - funkcją odwrotną przy $y \in [-1, +1]$ jest $x = \mathrm{acos}(y)$
        - » $\mathrm{acos}(y) \in [0, \pi]$

  więc liczby z przedziału $[-1,+1]$ mogą służyć jako miary kąta α

Miara kąta pomiędzy wektorami

- Twierdzenie Cauchy'go–Bunyakovskii'ego–Schwarz'a:
  - dla dowolnych wektorów zachodzi:

    $$|\mathbf{x}^T\mathbf{y}| \leq ||\mathbf{x}||||\mathbf{y}||$$

    gdzie
    - $|s|$ – wartość bezwzględna (skalara)
    - $||\mathbf{u}||$ – norma Euklidesa (wektora)
  - zależność $|\mathbf{x}^T\mathbf{y}| \leq ||\mathbf{x}||||\mathbf{y}||$ implikuje oczywiście zależność $\mathbf{x}^T\mathbf{y} \leq ||\mathbf{x}||||\mathbf{y}||$ (ponieważ $\mathbf{x}^T\mathbf{y} \leq |\mathbf{x}^T\mathbf{y}|$)
  - wniosek: $-1 \leq \mathbf{x}^T\mathbf{y}/(||\mathbf{x}||||\mathbf{y}||) \leq +1$

Miara kąta pomiędzy wektorami

- Wartość cosinusa kąta między niezerowymi wektorami **x** i **y**

  $\cos(\alpha) = \mathbf{x}^T\mathbf{y}/(||\mathbf{x}||||\mathbf{y}||) = (\mathbf{x}/||\mathbf{x}||)^T(\mathbf{y}/||\mathbf{y}||)$

- W szczególności
  - gdy $||\mathbf{x}|| = 1$ i $||\mathbf{y}|| = 1$, to $\cos(\alpha) = \mathbf{x}^T\mathbf{y}$
  - gdy **x** i **y** są współliniowe, tzn. $\mathbf{y} = a\mathbf{x}$, gdzie a jest niezerowym skalarem, to $\cos(\alpha) = \pm 1$, przy czym
    - $\cos(\alpha) = +1$ dla $a > 0$
    - $\cos(\alpha) = -1$ dla $a < 0$

…

Rozkład EVD macierzy

- Ze względu na parametryczność rozwiązań układu równań definiującego wektory własne, możliwe jest tworzenie bardzo różnych instancji tych wektorów
  - w wielu różnych zastosowaniach parametry dobiera się w taki sposób, aby powstałe wektory własne były:
    - unormowane: $(\mathbf{k}_i)^T\mathbf{k}_i = 1$
    - wzajemnie ortogonalne: $(\mathbf{k}_i)^T\mathbf{k}_j = 0$ dla $i \neq j$

    (czyli: tak, aby macierz wektorów własnych była ortogonalna)
  - rozwiązania tej postaci są także najczęściej generowane przez różne funkcje/biblioteki komputerowe służące do generowania wektorów własnych

Rozkład EVD macierzy symetrycznych

- Twierdzenie
  - dla macierzy symetrycznych wektory własne odpowiadające różnym wartościom własnym są wzajemnie ortogonalne
  - uzasadnienie wykorzystujące
    - właściwość macierzy symetrycznej $\mathbf{A}$: $\mathbf{A}^T=\mathbf{A}$
    - definicję wektora własnego macierzy $\mathbf{A}$: $\mathbf{Ak}=\lambda\mathbf{k}$
    - ogólne prawa, tzn. np.: $(\mathbf{AB})^T=\mathbf{B}^T\mathbf{A}^T$ oraz $\mathbf{A}=\mathbf{B}\Rightarrow\mathbf{A}^T=\mathbf{B}^T$

Rozkład EVD macierzy symetrycznych

- Niech dane będą dwie wartości własne λ_i oraz λ_j macierzy symetrycznej **A** takie, że $\lambda_i - \lambda_j \neq 0$

$\mathbf{A}\mathbf{k}_i = \lambda_i \mathbf{k}_i$ $\qquad$ $\mathbf{A}\mathbf{k}_j = \lambda_j \mathbf{k}_j$

$(\mathbf{k}_j)^T\mathbf{A}\mathbf{k}_i = (\mathbf{k}_j)^T\lambda_i\mathbf{k}_i$ (*1*) $\qquad$ $(\mathbf{k}_i)^T\mathbf{A}\mathbf{k}_j = (\mathbf{k}_i)^T\lambda_j\mathbf{k}_j$ (*2*)

$((\mathbf{k}_j)^T\mathbf{A}\mathbf{k}_i)^T = ((\mathbf{k}_j)^T\lambda_i\mathbf{k}_i)^T$

$(\mathbf{k}_i)^T\mathbf{A}^T\mathbf{k}_j = (\mathbf{k}_i)^T(\lambda_i)^T\mathbf{k}_j$

$(\mathbf{k}_i)^T\mathbf{A}\mathbf{k}_j = (\mathbf{k}_i)^T\lambda_i\mathbf{k}_j$ (*3*)

  – odejmując równania (*3*) oraz (*2*) stronami otrzymujemy

$(\mathbf{k}_i)^T\mathbf{A}\mathbf{k}_j - (\mathbf{k}_i)^T\mathbf{A}\mathbf{k}_j = (\mathbf{k}_i)^T\lambda_i\mathbf{k}_j - (\mathbf{k}_i)^T\lambda_j\mathbf{k}_j$

$0 = (\mathbf{k}_i)^T\lambda_i\mathbf{k}_j - (\mathbf{k}_i)^T\lambda_j\mathbf{k}_j$

$0 = \lambda_i(\mathbf{k}_i)^T\mathbf{k}_j - \lambda_j(\mathbf{k}_i)^T\mathbf{k}_j$

$0 = (\lambda_i - \lambda_j)(\mathbf{k}_i)^T\mathbf{k}_j$

  – a z tego wynika, że $(\mathbf{k}_i)^T\mathbf{k}_j=0$ (ponieważ $\lambda_i - \lambda_j \neq 0$)

Rozkład EVD macierzy symetrycznych

- Twierdzenie
  - dla macierzy symetrycznych można tak dobrać wektory własne odpowiadające identycznym wartościom własnym, aby były wzajemnie ortogonalne
  - uzasadnienie

    (pominięte)

Rozkład EVD macierzy symetrycznych

- Dla każdej symetrycznej macierzy **A** istnieją macierze: **K** i diagonalna **L** takie, że **AK** = **KL**
- Ale
  - jeżeli **A** jest symetryczna, to można tak dobrać jej wektory własne, że **K** jest macierzą ortogonalną
    - oznacza to, że $\mathbf{K}^{T} = \mathbf{K}^{-1}$
  - wtedy

    $\mathbf{AK} = \mathbf{KL}$

    $\mathbf{AKK}^{T} = \mathbf{KLK}^{T}$

    $\mathbf{AI} = \mathbf{KLK}^{T}$

    $\mathbf{A} = \mathbf{KLK}^{T}$
- Wniosek
  - jeżeli **A** jest symetryczna, to posiada rozkład EVD postaci $\mathbf{KLK}^{T}$

…