...

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

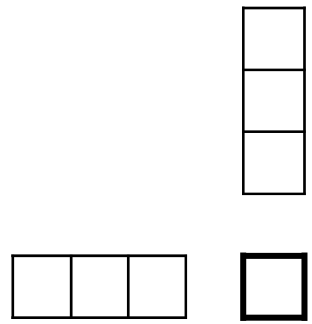
Prezentowane materiały, będące dodatkiem pomocniczym do wykładów, z konieczności fragmentarycznym i niedopracowanym, należy wykorzystywać z pełną świadomością faktu, że mogą nie być pozbawione przypadkowych błędów, braków, wypaczeń i przeinaczeń :-)

Autor

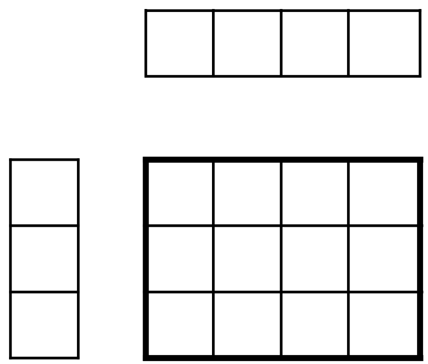
…

Iloczyny wektorów

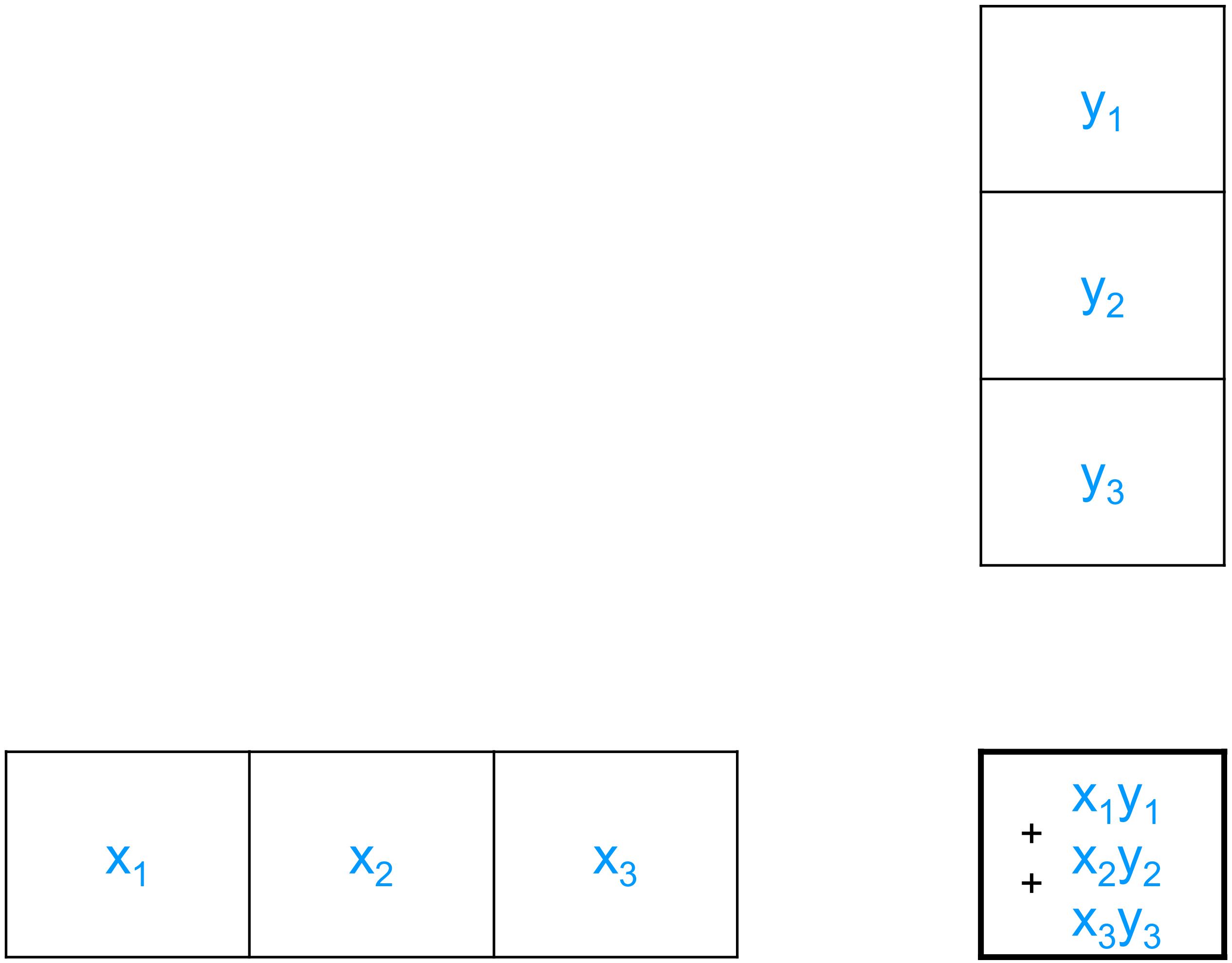
- Iloczyn skalarny wektorów:
  - wierszowy jest mnożony przez kolumnowy
  - długości wektorów muszą być równe
  - wynikiem jest skalar



- Iloczyn macierzowy wektorów:
  - kolumnowy jest mnożony przez wierszowy
  - długości wektorów nie muszą być równe
  - wynikiem jest macierz



„Mechanika” iloczynu skalarnego



„Mechanika” iloczynu macierzowego

x_1	x_2	x_3	x_4

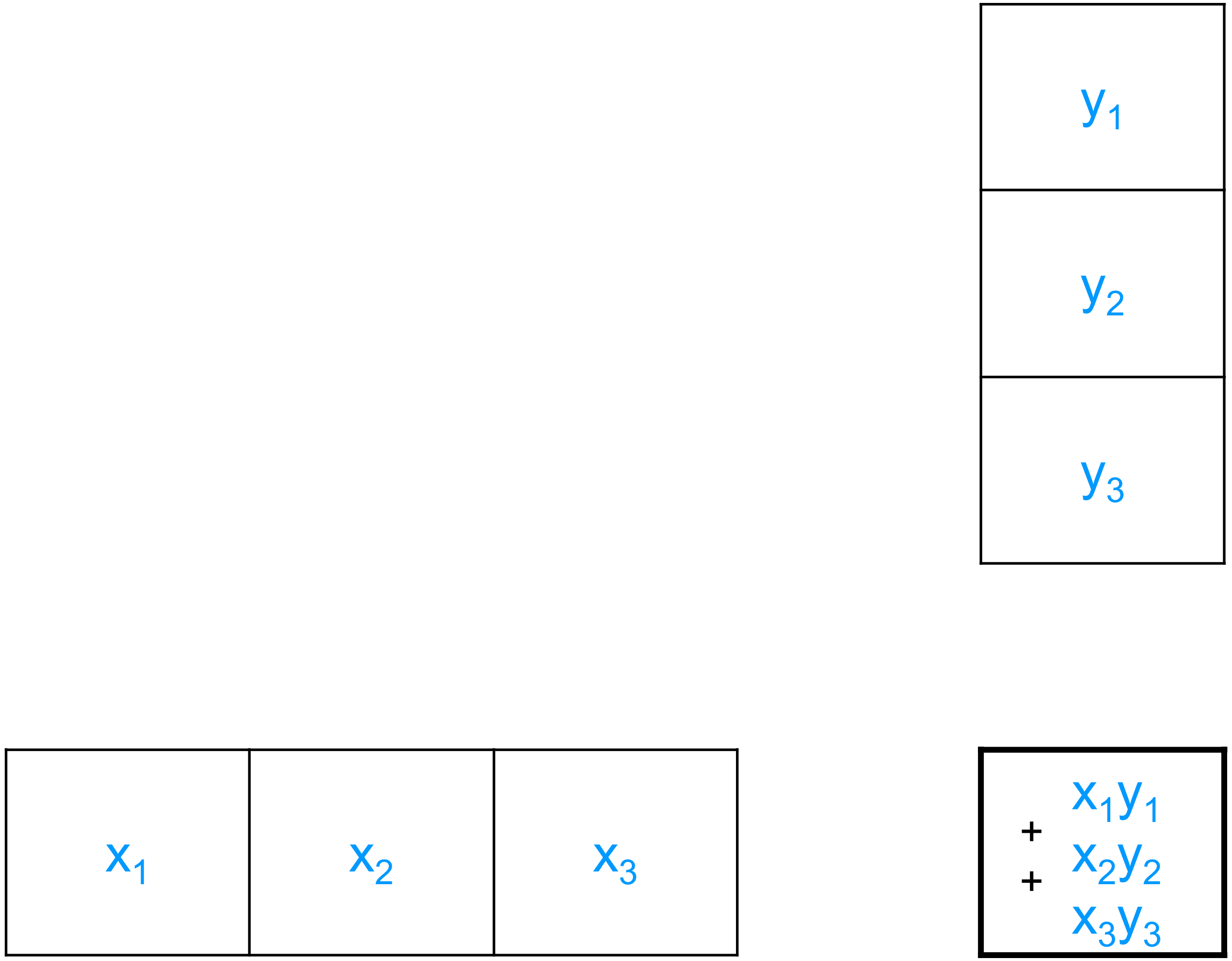
y_1
y_2
y_3

y_1x_1	y_1x_2	y_1x_3	y_1x_4
y_2x_1	y_2x_2	y_2x_3	y_2x_4
y_3x_1	y_3x_2	y_3x_3	y_3x_4

Iloczyn macierzy: złożenie iloczynów skalarnych wektorów

- Iloczyn macierzy może być przedstawiony jako:
  - iloczyn skalarny wektorów
    - każdy element macierzy wynikowej jest ilorazem skalarnym odpowiednich wektorów będących elementami mnożonych macierzy
  - iloczyn macierzowy wektorów
    - macierz wynikowa jest sumą pewnych macierzy (tzw. warstw) powstałych jako ilorazy macierzowe odpowiednich wektorów będących elementami mnożonych macierzy

Iloczyn macierzy:
złożenie iloczynów skalarnych wektorów



Iloczyn macierzy:
złożenie iloczynów skalarnych wektorów

y_{11}	y_{12}
y_{21}	y_{22}
y_{31}	y_{32}

x_{11}	x_{12}	x_{13}
x_{21}	x_{22}	x_{23}

$x_{11}y_{11} + x_{12}y_{21} + x_{13}y_{31}$	$x_{12}y_{12} + x_{12}y_{22} + x_{13}y_{32}$
$x_{21}y_{11} + x_{22}y_{21} + x_{23}y_{31}$	$x_{21}y_{12} + x_{22}y_{22} + x_{23}y_{32}$

Iloczyn macierzy: suma iloczynów macierzowych wektorów

	x_1	x_2	x_3	x_4
y_{12}	$y_{12}x_{21}$	$y_{12}x_{22}$	$y_{12}x_{23}$	$y_{12}x_{24}$
y_{22}	$y_{22}x_{21}$	$y_{22}x_{22}$	$y_{22}x_{23}$	$y_{22}x_{24}$
y_{32}	$y_{32}x_{21}$	$y_{32}x_{22}$	$y_{32}x_{23}$	$y_{32}x_{24}$

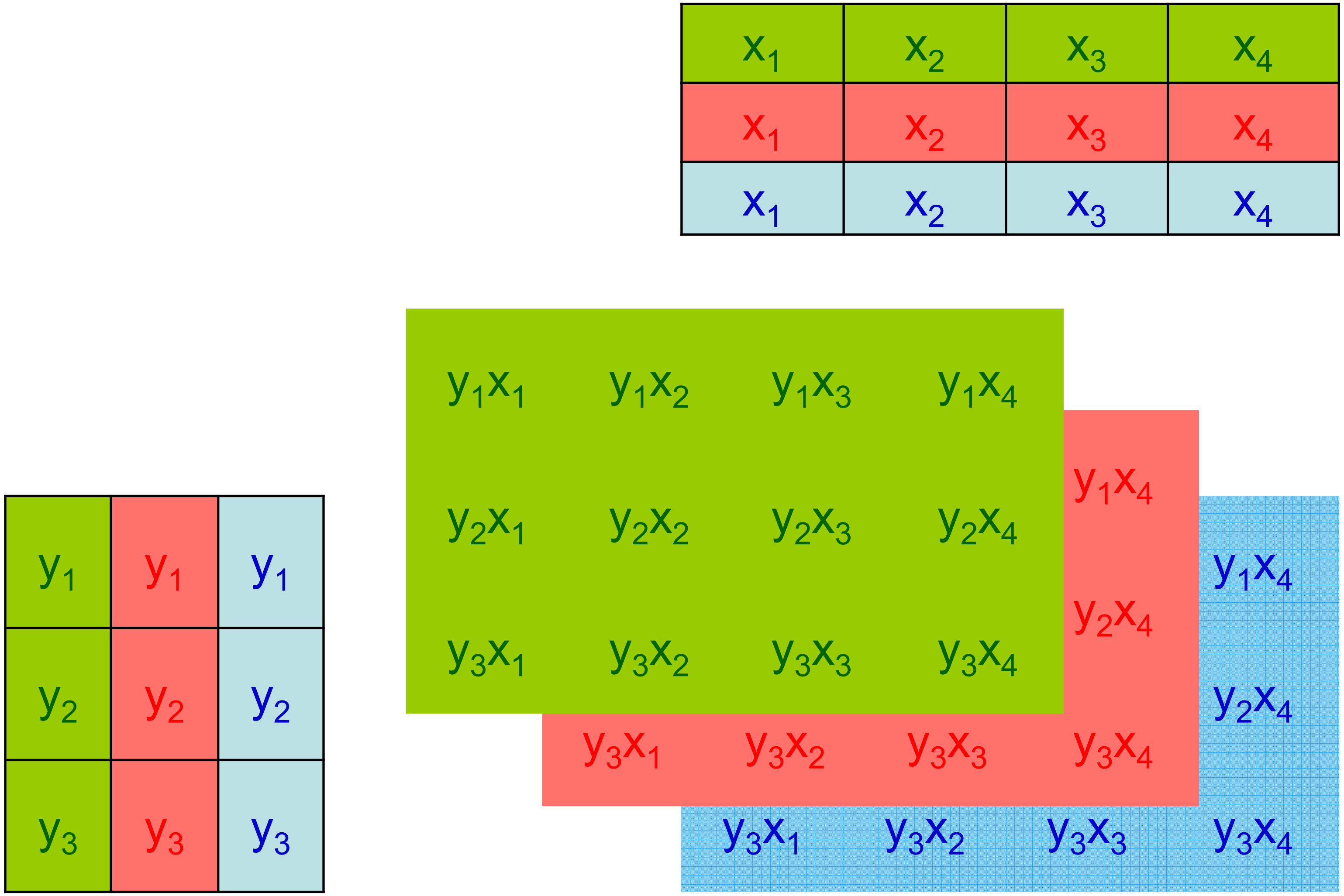
Iloczyn macierzy:
suma iloczynów macierzowych wektorów

x_{11}	x_{12}	x_{13}	x_{14}
x_{21}	x_{22}	x_{23}	x_{24}

y_{11}	y_{12}
y_{21}	y_{22}
y_{31}	y_{32}

$y_{11}x_{11}$ + $y_{12}x_{21}$	$y_{11}x_{12}$ + $y_{12}x_{22}$	$y_{11}x_{13}$ + $y_{12}x_{23}$	$y_{11}x_{14}$ + $y_{12}x_{24}$
$y_{21}x_{11}$ + $y_{22}x_{21}$	$y_{21}x_{12}$ + $y_{22}x_{22}$	$y_{21}x_{13}$ + $y_{22}x_{23}$	$y_{21}x_{14}$ + $y_{22}x_{24}$
$y_{31}x_{11}$ + $y_{32}x_{21}$	$y_{31}x_{12}$ + $y_{32}x_{22}$	$y_{31}x_{13}$ + $y_{32}x_{23}$	$y_{31}x_{14}$ + $y_{32}x_{24}$

Iloczyn macierzy:
suma iloczynów macierzowych wektorów



Iloczyn „skalowany”

- O „skalowanym” iloczynie macierzy **X** i **Z** można mówić w sytuacji sytuacji, gdy wykonywana jest operacja **XDZ**, gdzie **D** jest macierzą diagonalną

  – Dla **W** = **XZ** (iloczyn „nieskalowany”)

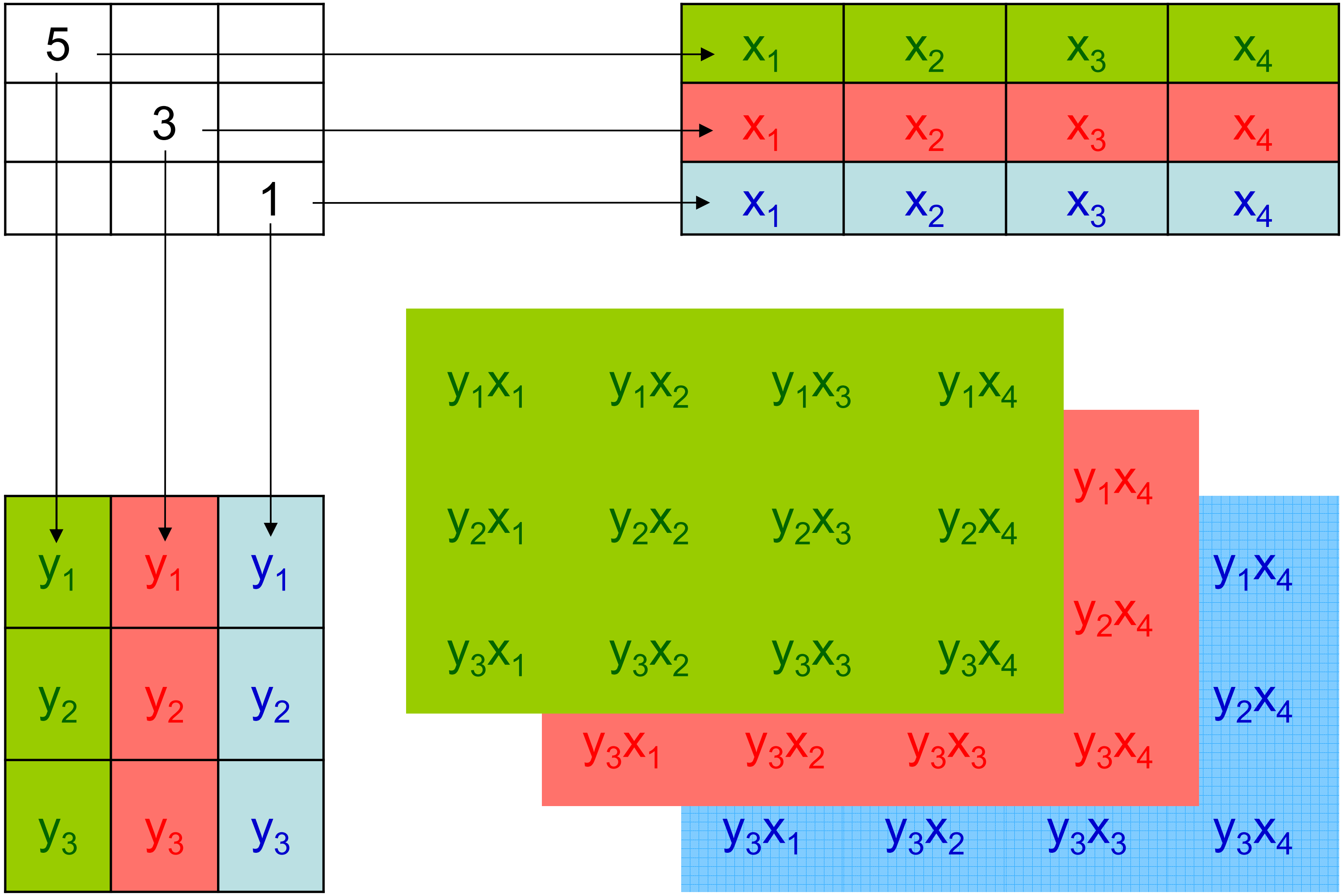
    - zachodzi: $\mathrm{w}_{\mathrm{kl}} = \sum_{i} \mathrm{x}_{\mathrm{k}i} \cdot \mathrm{z}_{i1}$

  – Dla **W** = **XDZ** (iloczyn „skalowany”)

    - zachodzi:

$$\mathrm{w}_{kl} = \sum_{i} \mathrm{x}_{ki} \cdot d_{ii} \cdot \mathrm{z}_{il} = \sum_{i} d_{ii} \cdot x_{ki} \cdot \mathrm{z}_{il}$$

„Mechanika” skalowanego iloczynu macierzy

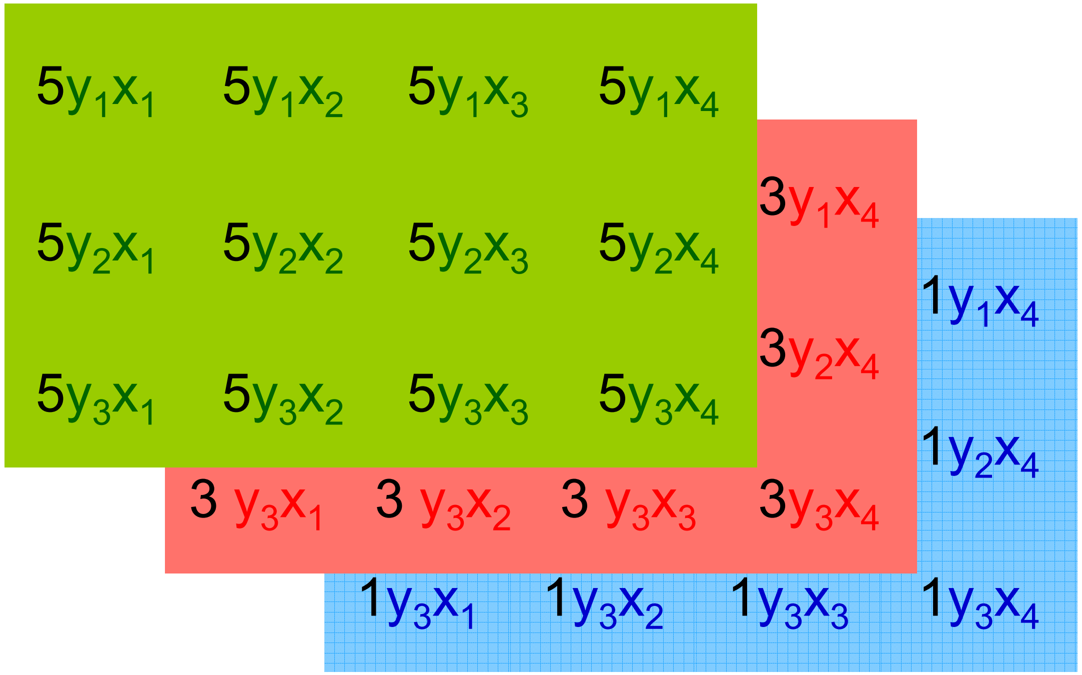


Skalowany iloczyn macierzy

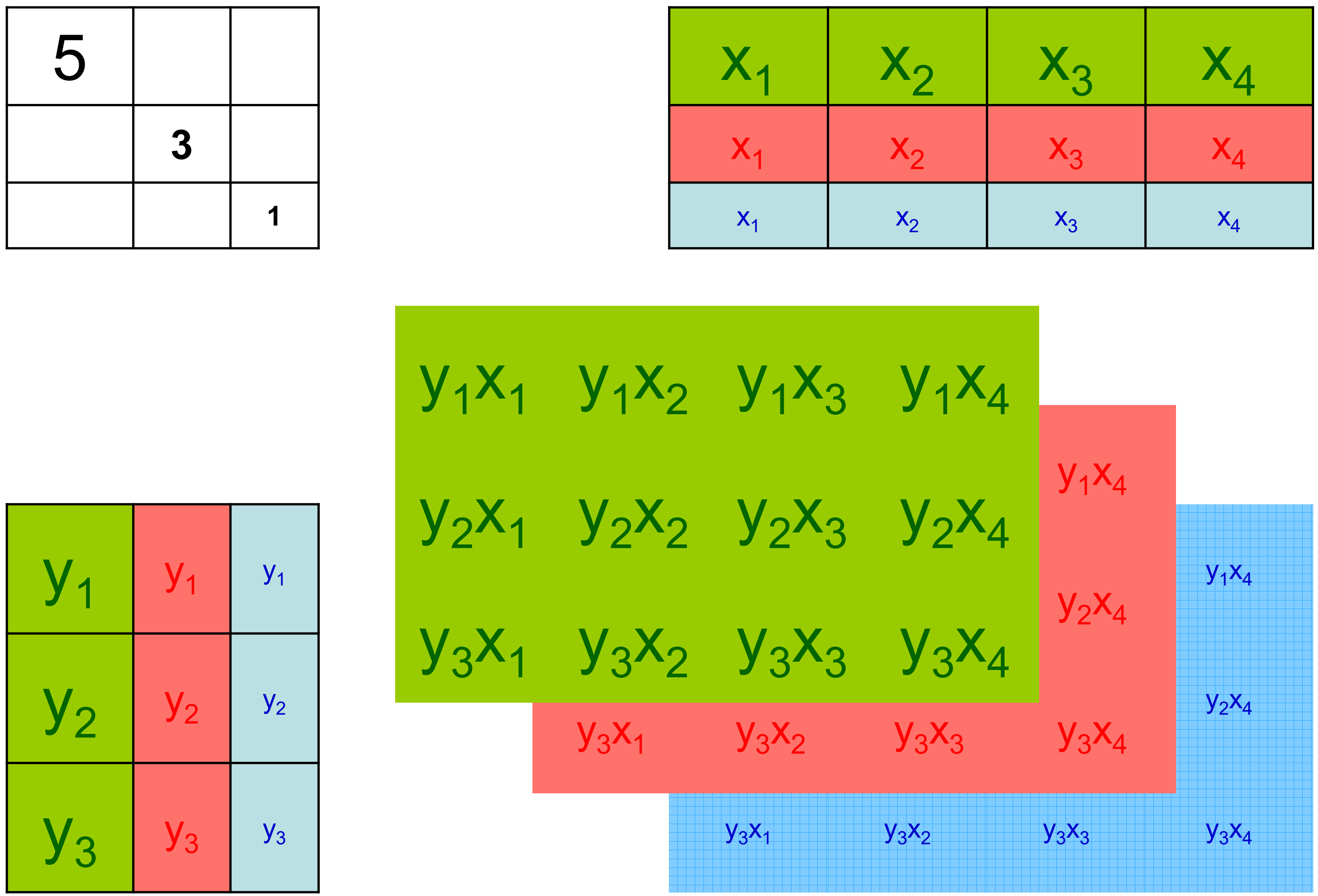
5		
	3	
		1

x_1	x_2	x_3	x_4
x_1	x_2	x_3	x_4
x_1	x_2	x_3	x_4

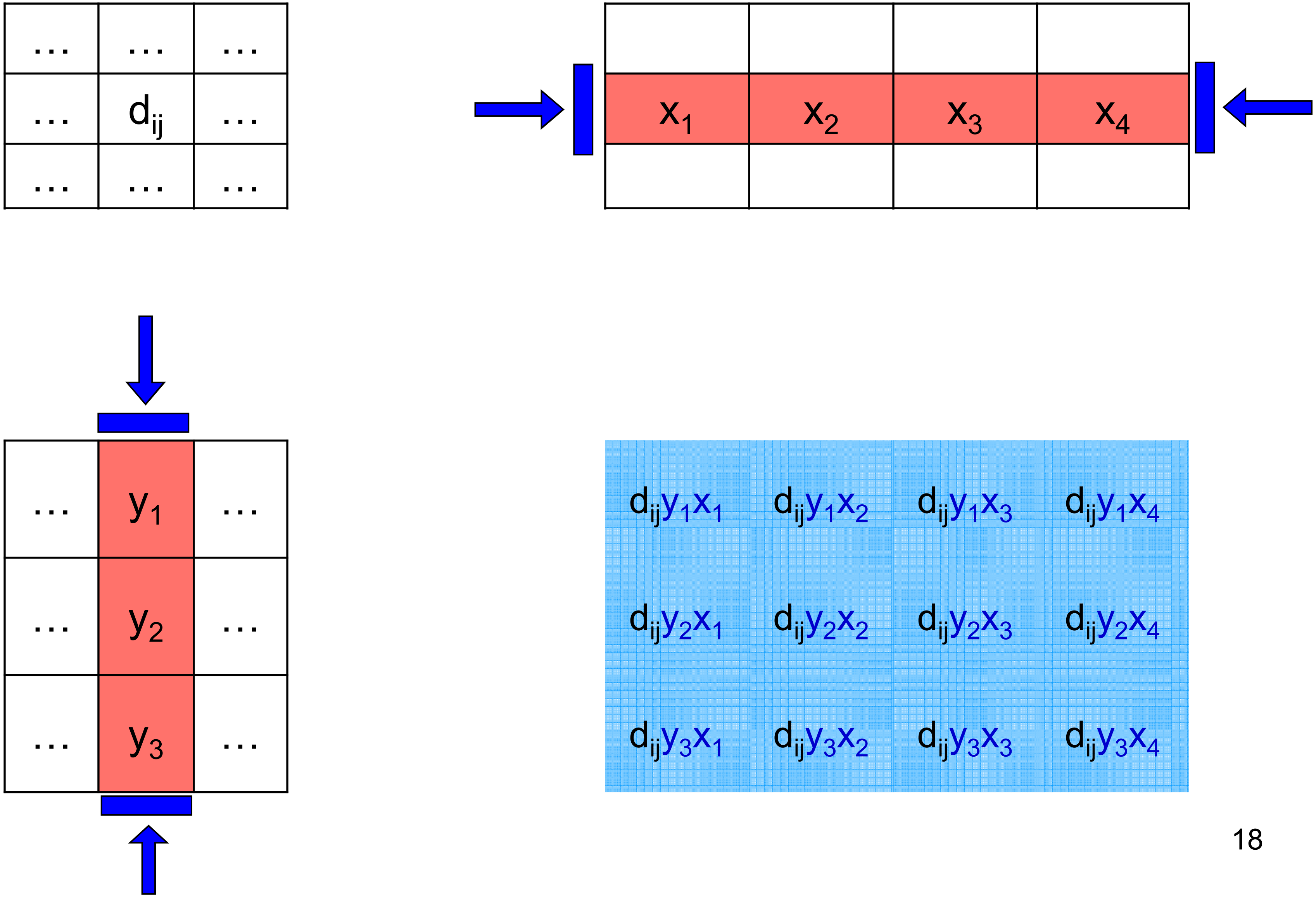
y_1	y_1	y_1
y_2	y_2	y_2
y_3	y_3	y_3



Skalowany iloczyn macierzy



Iloczyn trzech macierzy jako iloczyn skalowany



Iloczyn „skalowany”

- Przypadek skalowania
  - lewostronnego
    - $\mathbf{W} = (\mathbf{XD})\mathbf{Z}$ (macierzą **D** skalujemy macierz po lewej)
  - prawostronnego
    - $\mathbf{W} = \mathbf{X}(\mathbf{DZ})$ (macierzą **D** skalujemy macierz po prawej)
  - „równomiernego” (jednocześnie lewo- i prawostronnego)
    - $\mathbf{W} = (\mathbf{XD}^{0.5})(\mathbf{D}^{0.5}\mathbf{Z})$

    - przykładowe zastosowanie: macierz $\mathbf{X} = \mathbf{KL}^{0.5}$ w MDS
      - wtedy **X** jest (poszukiwaną) macierzą spełniającą $\mathbf{XX}^T = \mathbf{B}$, ponieważ

$$\mathbf{XX}^T = \mathbf{KL}^{0.5}(\mathbf{KL}^{0.5})^T = \mathbf{KL}^{0.5}(\mathbf{L}^{0.5})^T\mathbf{K}^T =$$
$$= \mathbf{KL}^{0.5}\mathbf{L}^{0.5}\mathbf{K}^T = \mathbf{KL}^1\mathbf{K}^T = \mathbf{KLK}^T = \mathbf{B}$$

…

Przybliżenia macierzy dzięki rozkładowi EVD

- Niech **A** będzie macierzą danych, oraz niech istnieją macierze **X**, **D** i **Z** takie, że: **A** = **XDZ**, przy czym **D** jest macierzą diagonalną o wartościach d_{ii} na przekątnej:
  - jeżeli pewne wartości d_{ii} są co do wartości bezwzględnej małe w porównaniu z innymi d_{ii}, to elementy te wraz z:
    - odpowiadającymi im kolumnami macierzy **X**
    - odpowiadającymi im wierszami macierzy **Z**

    mogą zostać pominięte (uznane za równe zero)
  - otrzymany w rezultacie takiej operacji, uproszczony iloczyn macierzowy będzie przybliżeniem (lepszym lub gorszym) oryginalnej macierzy **A**
- Stworzenie takiego przybliżenia może być celem samym w sobie, ponieważ może doprowadzić do m.in.:
  - „odszumienia” obrazu reprezentowanego przez macierz **A**
  - zredukowanie reprezentacji pamięciowej macierzy **A**
- W praktyce operacja taka jest możliwa dzięki rozkładowi **A** = **KLK**^{T}, w którym macierz **L** z definicji jest macierzą diagonalną

…

Uogólnienie EVD: rozkład SVD

- Niech dana będzie macierz symetryczna $\mathbf{A}$ oraz jej rozkład EVD postaci $\mathbf{A} = \mathbf{K}\mathbf{L}\mathbf{K}^T$
- Pytanie, czy rozkład powyższej postaci, lub jakieś jego uogólnienie, może być zagwarantowany/e dla macierzy, które:
  - nie są symetryczne?
  - nie są kwadratowe?
  - nie są rzeczywiste?

Uogólnienie EVD: rozkład SVD

- Dana jest macierz **A** o rozmiarach mxn
  - jeżeli m = n, to niektóre **A** można rozłożyć na iloczyn czynników $\mathbf{KLK}^{-1}$ (o ile $\mathbf{K}^{-1}$ istnieje, jeżeli tak to w szczególnym przypadku może zachodzić: $\mathbf{K}^{-1} = \mathbf{K}^{T}$) wykorzystując do tego macierze złożone z wartości własnych (**L**) i wektorów własnych (**K**) macierzy **A**
  - pytanie, czy jeżeli $m \neq n$ to także można wykorzystać jakieś wartości i wektory własne do znalezienia takiego rozkładu?
  - jak znaleźć te wartości jeżeli wiadomo, że wartości/wektory własne oblicza się jedynie dla macierzy kwadratowych?
- W celu uogólnienia rozkładu macierzy niekwadratowej na iloczyn trzech czynników wykorzystuje się następujący fakt:
  - nawet jeżeli **A** nie jest macierzą symetryczną ani kwadratową (czyli gdy $m \neq n$) to można wykorzystać tę macierz do utworzenia dwóch macierzy:
    - $\mathbf{C} = \mathbf{A}^{T}\mathbf{A}$ (odpowiednik macierzy kowariancji kolumn), o rozmiarach nxn
    - $\mathbf{R} = \mathbf{A}\mathbf{A}^{T}$ (odpowiednik macierzy kowariancji wierszy), o rozmiarach mxm

Uogólnienie EVD: rozkład SVD

- Macierze utworzone z dowolnej macierzy $\mathbf{A}$ poprzez wykonanie operacji $\mathbf{A}^T\mathbf{A}$ oraz $\mathbf{A}\mathbf{A}^T$ nazywa się macierzami Grama
  - ogólnie macierz Grama tworzy się stosując formę (czyli funkcję, której argumentami są wektory) postaci f: $R^n x R^n \rightarrow R$ (np. $f(\mathbf{x},\mathbf{y}) = \mathbf{x}^T\mathbf{A}\mathbf{y}$) do każdej pary wektorów ze zbioru $[\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_m]$, w rezultacie czego powstaje macierz (składająca się z wartości formy) o rozmiarach mxm
    - iloczyn skalarny wektorów, który jest wykorzystany do stworzenia macierzy $\mathbf{A}^T\mathbf{A}$ oraz $\mathbf{A}\mathbf{A}^T$ jest szczególnym przypadkiem takiej formy: $\langle\mathbf{x},\mathbf{y}\rangle = \mathbf{x}^T\mathbf{y} = \mathbf{x}^T\mathbf{I}\mathbf{y}$
    - przekształceniu przez tę formę podlegają kolumny/wiersze macierzy $\mathbf{A}$
  - jeżeli forma $f(\mathbf{x},\mathbf{y})$ jest dodatnio określona to stworzona macierz Grama ma wyznacznik nieujemny
    - iloczyn skalarny wektorów spełnia warunek dodatniej określoności
  - wyznacznik macierzy Grama utworzonej dla wektorów $\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_m$ nie zależy od kolejności tych wektorów
    - tzn. np. $\det([\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_m]) = \det([\mathbf{w}_2, \mathbf{w}_1, ..., \mathbf{w}_m])$
    - uzasadnienie: macierz Grama utworzoną dla $[\mathbf{w}_2, \mathbf{w}_1, ..., \mathbf{w}_m]$ można utworzyć z macierzy Grama utworzonej dla $[\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_m]$ przez zamianę miejscami pierwszych dwóch wierszy i pierwszych dwóch kolumn

Uogólnienie EVD: rozkład SVD

- Wartości wyznaczników obu macierzy Grama (utworzonych jako $\mathbf{C} = \mathbf{A}^T\mathbf{A}$ oraz $\mathbf{R} = \mathbf{A}\mathbf{A}^T$) są uzależnione od rzędu macierzy **A** (który świadczy o zależności/niezależności kolumn/wierszy tej macierzy)
  - jeżeli kolumny macierzy **A** są liniowo niezależne to $\det(\mathbf{C}) = \det(\mathbf{A}^T\mathbf{A}) > 0$, w przeciwnym przypadku $\det(\mathbf{C}) = \det(\mathbf{A}^T\mathbf{A}) = 0$
  - jeżeli wiersze macierzy **A** są liniowo niezależne to $\det(\mathbf{R}) = \det(\mathbf{A}\mathbf{A}^T) > 0$, w przeciwnym przypadku $\det(\mathbf{R}) = \det(\mathbf{A}\mathbf{A}^T) = 0$
- Wartości własne macierzy Grama są nieujemne
  - wynika to z faktu, macierze grama są kwadratowymi macierzami symetrycznymi

Uogólnienie EVD: rozkład SVD

- Szczególne przypadki
  - dla kwadratowej macierzy **A** wyznacznik macierzy $\mathbf{A}^T\mathbf{A}$ jest kwadratem wyznacznika macierzy **A** (co determinuje jego nieujemność)
    - $\det(\mathbf{A}^T\mathbf{A}) = \det(\mathbf{A}^T)\det(\mathbf{A}) = \det(\mathbf{A})\det(\mathbf{A}) = (\det(\mathbf{A}))^2$
  - za bardzo szczególny przypadek macierzy Grama można uznać macierz jednostkową **I**, która jest wynikiem operacji $\mathbf{K}^T\mathbf{K}$, gdzie **K** jest macierzą ortogonalną
    - oczywiście dla obu macierzy Grama utworzonych dla macierzy ortogonalnej **K** zachodzi:
      - $\det(\mathbf{C}) = \det(\mathbf{K}^T\mathbf{K}) = \det(\mathbf{I}) = 1 > 0$
      - $\det(\mathbf{R}) = \det(\mathbf{K}\mathbf{K}^T) = \det(\mathbf{I}) = 1 > 0$

Uogólnienie EVD: rozkład SVD

- Pomimo iż $\mathbf{C} = \mathbf{A}^T\mathbf{A}$ i $\mathbf{R} = \mathbf{A}\mathbf{A}^T$ są różnymi macierzami to posiadają one wiele cech wspólnych, m.in.:
  - $\mathbf{C}^T = \mathbf{C}$: **C** jest kwadratową macierzą symetryczną
  - $\mathbf{R}^T = \mathbf{R}$: **R** jest kwadratową macierzą symetryczną
  - spec(**C**) ≥ 0: wartości własne macierzy **C** są nieujemne
  - spec(**R**) ≥ 0: wartości własne macierzy **R** są nieujemne
  - rank(**C**) = rank(**R**): rzędy macierzy **C** i **R** są takie same
  - trace(**C**) = trace(**R**): ślady macierzy **C** i **R** są takie same
  - $\mathrm{spec}_{\neq 0}(\mathbf{C}) = \mathrm{spec}_{\neq 0}(\mathbf{R})$: zbiory niezerowych wartości własnych macierzy **C** i **R** są takie same
  - zbiory wektorów własnych obu macierzy są związane ze sobą pewną zależnością liniową

Uogólnienie EVD: rozkład SVD

- Inne cechy macierzy $\mathbf{C} = \mathbf{A}^T\mathbf{A}$ i $\mathbf{R} = \mathbf{A}\mathbf{A}^T$ to:
  - jeżeli m < n to:
    - $\det(\mathbf{C}) = 0$
    - $\mathrm{rank}(\mathbf{C}) < n$
  - jeżeli m > n to:
    - $\det(\mathbf{R}) = 0$
    - $\mathrm{rank}(\mathbf{R}) < m$

Uogólnienie EVD: rozkład SVD

- Wykorzystując EVD macierzy $\mathbf{C} = \mathbf{A}^T\mathbf{A}$ i $\mathbf{R} = \mathbf{A}\mathbf{A}^T$ można zapisać zależności:

  $\mathbf{C} = \mathbf{V}\mathbf{L}_V\mathbf{V}^{-1}$ i $\mathbf{R} = \mathbf{U}\mathbf{L}_U\mathbf{U}^{-1}$

  gdzie:
  - $\mathbf{V}$ i $\mathbf{L}_V$ są macierzami (odpowiednio) wektorów i wartości własnych macierzy $\mathbf{C}$
    - $\mathbf{C}$ jest symetryczna, a więc $\mathbf{V}^{-1}$ istnieje, i dodatkowo: $\mathbf{V}^{-1} = \mathbf{V}^T$
  - $\mathbf{U}$ i $\mathbf{L}_U$ są macierzami (odpowiednio) wektorów i wartości własnych macierzy $\mathbf{R}$
    - $\mathbf{R}$ jest symetryczna, a więc $\mathbf{U}^{-1}$ istnieje, i dodatkowo: $\mathbf{U}^{-1} = \mathbf{U}^T$

  czyli

  $\mathbf{C} = \mathbf{V}\mathbf{L}_V\mathbf{V}^T$ i $\mathbf{R} = \mathbf{U}\mathbf{L}_U\mathbf{U}^T$
- Ostatecznie można zapisać:

  $\mathbf{A}^T\mathbf{A} = \mathbf{V}\mathbf{L}_V\mathbf{V}^T$ oraz $\mathbf{A}\mathbf{A}^T = \mathbf{U}\mathbf{L}_U\mathbf{U}^T$

Uogólnienie EVD: rozkład SVD

- Macierze **C** i **R** charakteryzują się identycznymi zbiorami niezerowych wartości własnych:

  $\mathrm{spec}_{\neq 0}(\mathbf{C}) = \mathrm{spec}_{\neq 0}(\mathbf{R})$
- Załóżmy, że
  - m > n, wtedy:
    - $\mathrm{spec}(\mathbf{C}) \subset \mathrm{spec}(\mathbf{R})$ – zbiór wartości własnych macierzy **R** zawiera wszystkie wartości własne macierzy **C** oraz m–n dodatkowych wartości zerowych, czyli:
      - $\mathrm{spec}(\mathbf{C}) \cup \{0,0,...,0\} = \mathrm{spec}(\mathbf{R})$
  - oczywiście dla m < n zachodzi:
    - $\mathrm{spec}(\mathbf{R}) \subset \mathrm{spec}(\mathbf{C})$
    - $\mathrm{spec}(\mathbf{R}) \cup \{0,0,...,0\} = \mathrm{spec}(\mathbf{C})$

Uogólnienie EVD: rozkład SVD

- Zależności pomiędzy zbiorami wartości własnych macierzy **C** i **R** przekładają się na zależności pomiędzy macierzami $\mathbf{L}_V$ i $\mathbf{L}_U$
- Dla m > n mamy:
  - $\mathbf{L}_V = \mathrm{diag}(\mathrm{spec}(\mathbf{C}))$
    - przekątna macierzy $\mathbf{L}_V$ (o rozmiarach nxn) zawiera n wartości własnych macierzy **C**
  - $\mathbf{L}_U = \mathrm{diag}(\mathrm{spec}(\mathbf{C}) \cup \{0,0,...,0\})$
    - przekątna macierzy $\mathbf{L}_U$ (o rozmiarach mxm) zawiera n wartości własnych macierzy **C** oraz m–n wartości zerowych
- Dla m < n mamy:
  - $\mathbf{L}_U = \mathrm{diag}(\mathrm{spec}(\mathbf{R}))$
    - przekątna macierzy $\mathbf{L}_U$ (o rozmiarach mxm) zawiera m wartości własnych macierzy **C**
  - $\mathbf{L}_V = \mathrm{diag}(\mathrm{spec}(\mathbf{R}) \cup \{0,0,...,0\})$
    - przekątna macierzy $\mathbf{L}_V$ (o rozmiarach nxn) zawiera m wartości własnych macierzy **C** oraz n–m wartości zerowych

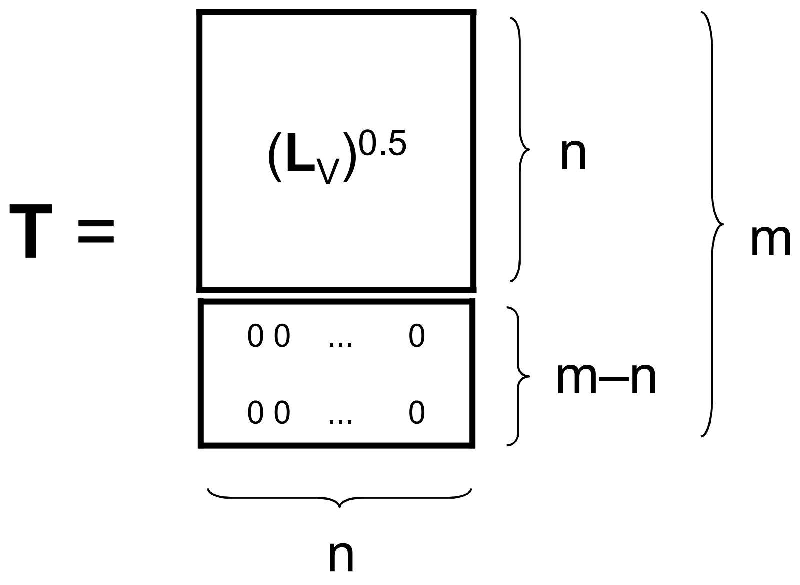
Uogólnienie EVD: rozkład SVD

- Gdy m > n to macierz $\mathbf{L}_U$ (o rozmiarach mxn) można przedstawić jako macierz $\mathbf{L}_V$ (o rozmiarach nxn) rozszerzoną o m–n kolumn zerowych i m–n wierszy zerowych:
  – zarówno $\mathbf{L}_U$ jaki i $\mathbf{L}_V$ pozostają macierzami diagonalnymi

$$\mathbf{L}_U = \begin{array}{cc} \left[\begin{array}{c|c} \mathbf{L}_V & \begin{matrix}0 & 0\\ 0 & 0\\ \dots & \\ 0 & 0\end{matrix} \\ \hline \begin{matrix}0\;0 & \dots & 0\\ 0\;0 & \dots & 0\end{matrix} & \begin{matrix}0 & \\ \dots & \\ & 0\end{matrix}\end{array}\right] & \left.\begin{matrix} \}\, n \\ \}\, m-n\end{matrix}\right\} m \\ \underbrace{\quad n \quad}\;\underbrace{m-n}_{} & \\ \underbrace{\qquad\qquad\quad}_{m} & \end{array}$$

Uogólnienie EVD: rozkład SVD

- Na podstawie macierzy $\mathbf{L}_U$ (lub $\mathbf{L}_V$) tworzy się nową macierz $\mathbf{T}$
- Dla m > n macierz $\mathbf{T}$ (o rozmiarach mxn) może być utworzona jako:
  - macierz $(\mathbf{L}_U)^{0.5}$, z której usunięto m–n prawych kolumn (zerowych) lub
  - macierz $(\mathbf{L}_V)^{0.5}$, do której dodano m–n dolnych kolumn (zerowych)



Uogólnienie EVD: rozkład SVD

- Diagonalne elementy macierzy **T** oblicza się jako pierwiastki wartości własnych obliczonych dla macierzy $\mathbf{A}^T\mathbf{A}$ lub $\mathbf{A}\mathbf{A}^T$
  - diagonalne elementy macierzy **T** są więc zdeterminowane przez macierze $\mathbf{A}^T\mathbf{A}$ lub $\mathbf{A}\mathbf{A}^T$ (są pierwiastkami ich wartości własnych)
  - ponieważ jednak $\mathbf{A}^T\mathbf{A}$ lub $\mathbf{A}\mathbf{A}^T$ powstają z macierzy **A**, można powiedzieć, że diagonalne elementy macierzy **T** są zdeterminowane przez macierz **A** – nazywamy je wartościami osobliwymi tej macierzy
- Dzięki specjalnej budowie macierzy **T** (o rozmiarach mxn) zachodzą związki:

  $\mathbf{T}^T\mathbf{T} = \mathbf{L}_V$ (rozmiar nxn)

  $\mathbf{T}\mathbf{T}^T = \mathbf{L}_U$ (rozmiar mxm)

Uogólnienie EVD: rozkład SVD

- Tworzymy iloczyn: $\mathbf{B} = \mathbf{U}\mathbf{T}\mathbf{V}^T$ z macierzy

  U: rozmiar mxm

  T: rozmiar mxn

  V: rozmiar nxn

/

Uogólnienie EVD: rozkład SVD

- Tworzymy iloczyn: $\mathbf{B} = \mathbf{U}\mathbf{T}\mathbf{V}^T$ z macierzy

  $\mathbf{U}$: rozmiar mxm

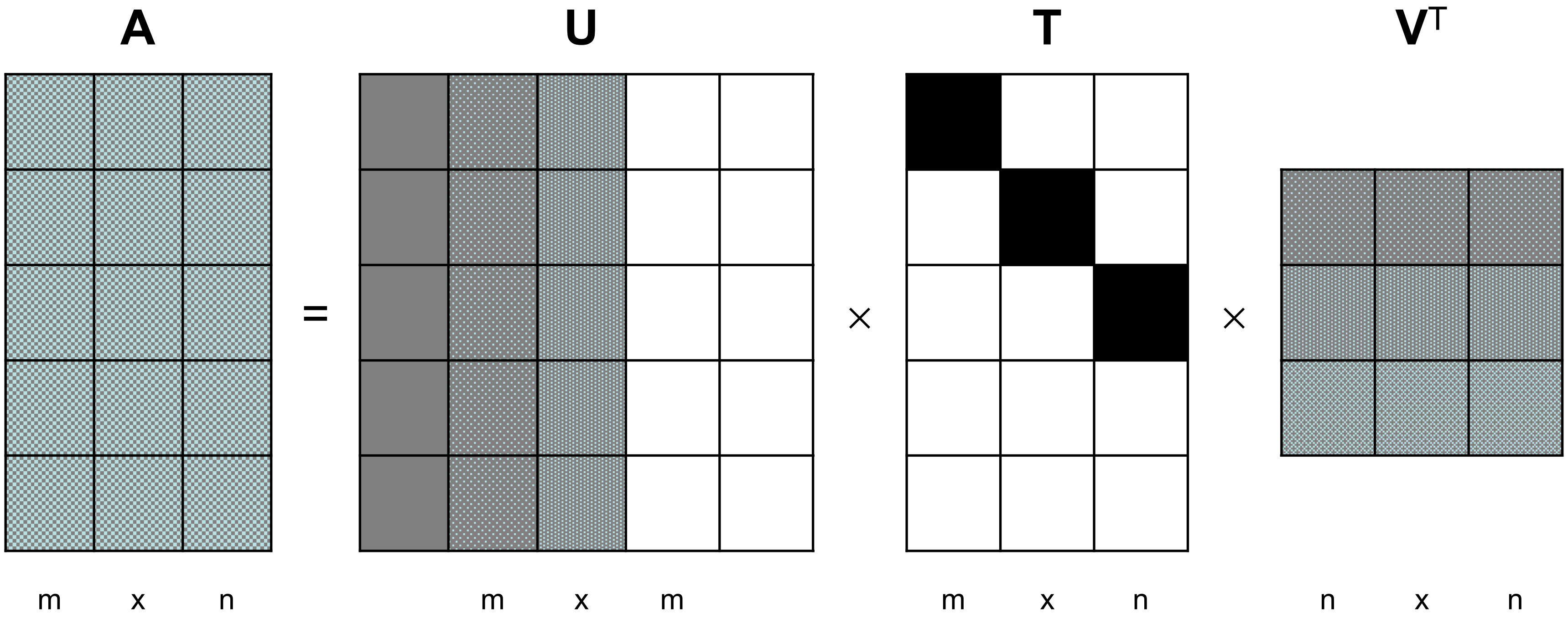
  $\mathbf{T}$: rozmiar mxn

  $\mathbf{V}$: rozmiar nxn
- Iloczyn ten można wykorzystać dalej do obliczenia: $\mathbf{B}^T\mathbf{B}$ oraz $\mathbf{B}\mathbf{B}^T$:

  $\mathbf{B}^T\mathbf{B} = (\mathbf{U}\mathbf{T}\mathbf{V}^T)^T(\mathbf{U}\mathbf{T}\mathbf{V}^T) = \mathbf{V}\mathbf{T}^T\mathbf{U}^T\mathbf{U}\mathbf{T}\mathbf{V}^T = \mathbf{V}\mathbf{T}^T\mathbf{I}\mathbf{T}\mathbf{V}^T = \mathbf{V}\mathbf{T}^T\mathbf{T}\mathbf{V}^T = \mathbf{V}\mathbf{L}_V\mathbf{V}^T$
  - wynika z tego, że $\mathbf{B}^T\mathbf{B} = \mathbf{A}^T\mathbf{A}$

  $\mathbf{B}\mathbf{B}^T = (\mathbf{U}\mathbf{T}\mathbf{V}^T)(\mathbf{U}\mathbf{T}\mathbf{V}^T)^T = \mathbf{U}\mathbf{T}\mathbf{V}^T\mathbf{V}\mathbf{T}^T\mathbf{U}^T = \mathbf{U}\mathbf{T}\mathbf{I}\mathbf{T}^T\mathbf{U}^T = \mathbf{U}\mathbf{T}\mathbf{T}^T\mathbf{U}^T = \mathbf{U}\mathbf{L}_U\mathbf{U}^T$
  - wynika z tego, że $\mathbf{B}\mathbf{B}^T = \mathbf{A}\mathbf{A}^T$
- Na tym etapie można podejrzewać, że $\mathbf{B} = \mathbf{A}$, czyli że wyrażenie postaci $\mathbf{U}\mathbf{T}\mathbf{V}^T$ jest poszukiwanym rozkładem macierzy $\mathbf{A}$ (bo $\mathbf{U}\mathbf{T}\mathbf{V}^T = \mathbf{B} = \mathbf{A}$)
  - ale: $a = b \Rightarrow a^2 = b^2$, jednak $a^2 = b^2 \not\Rightarrow a = b$
  - a więc równości $\mathbf{B}^T\mathbf{B} = \mathbf{A}^T\mathbf{A}$ oraz $\mathbf{B}\mathbf{B}^T = \mathbf{A}\mathbf{A}^T$ nie implikują $\mathbf{A} = \mathbf{B}$

„Mechanika” rozkładu SVD



Uogólnienie EVD: rozkład SVD

- Podsumowanie:
  - iloczyn $\mathbf{UTV}^T$ wyraża macierz $\mathbf{B}$ taką, że $\mathbf{B}^T\mathbf{B} = \mathbf{A}^T\mathbf{A}$ oraz $\mathbf{BB}^T = \mathbf{AA}^T$
  - i chociaż nie wynika z tego natychmiast fakt, że $\mathbf{B} = \mathbf{A}$, to w praktyce możliwe jest takie dobranie macierzy $\mathbf{U}$ i $\mathbf{V}$ aby zachodziło $\mathbf{B} = \mathbf{UTV}^T = \mathbf{A}$
- Wniosek:
  - dowolną macierz $\mathbf{A}$ można przedstawić w postaci iloczynu macierzy $\mathbf{UTV}^T$, które są utworzone z odpowiednio dobranych wektorów i wartości własnych macierzy $\mathbf{A}^T\mathbf{A}$ oraz $\mathbf{AA}^T$

Uogólnienie EVD: rozkład SVD

- Rozkład $\mathbf{A} = \mathbf{UTV}^T$ nazywa się rozkładem macierzy na wartości osobliwe (ang. Singular Value Decomposition, SVD)
- Przedstawiony sposób uzyskania rozkładu SVD macierzy $\mathbf{A}$ poprzez dokonanie rozkładów EVD macierzy $\mathbf{A}^T\mathbf{A}$ oraz $\mathbf{AA}^T$ nie jest stosowany w praktyce
  - problemy stanowią:
    - konieczność dobierania macierzy **U** i **V**
    - stabilność numeryczna

Przykład rozkładu SVD

- Symetryczna macierz kwadratowa 2x2

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}$$

$$\mathbf{A} = \underbrace{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}}_{\mathbf{U}} * \underbrace{\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}}_{\mathbf{T}} * \underbrace{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}}_{\mathbf{V}^T}$$

Przykład rozkładu SVD

$$\mathbf{C} = \underset{\mathbf{A}^T}{\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}} * \underset{\mathbf{A}}{\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$$

$$\mathbf{C} = \underset{\mathbf{K}}{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}} * \underset{\mathbf{L}}{\begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix}} * \underset{\mathbf{K}^T}{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}}$$

$$\mathbf{R} = \underset{\mathbf{A}}{\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}} * \underset{\mathbf{A}^T}{\begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}} = \begin{bmatrix} 5 & 4 \\ 4 & 5 \end{bmatrix}$$

$$\mathbf{R} = \underset{\mathbf{K}}{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}} * \underset{\mathbf{L}}{\begin{bmatrix} 9 & 0 \\ 0 & 1 \end{bmatrix}} * \underset{\mathbf{K}^T}{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}}$$

Przykład rozkładu SVD

T= $\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}$

U= $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

V= $\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

$$\underbrace{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}}_{\mathbf{U}} * \underbrace{\begin{bmatrix} 3 & 0 \\ 0 & 1 \end{bmatrix}}_{\mathbf{T}} * \underbrace{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ -1 & 1 \end{bmatrix}}_{\mathbf{V}^T} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} = \mathbf{A}$$

Przykład rozkładu SVD

- Niesymetryczna macierz kwadratowa 2x2

$$\mathbf{A}= \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}$$

$$\mathbf{A}= \underbrace{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{\mathbf{U}} * \underbrace{\begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix}}_{\mathbf{T}} * \underbrace{\frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}}_{\mathbf{V}^T}$$

Przykład rozkładu SVD

$$\mathbf{C} = \overset{\mathbf{A}^T}{\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}} * \overset{\mathbf{A}}{\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}} = \begin{bmatrix} 8 & 4 \\ 4 & 2 \end{bmatrix}$$

$$\mathbf{C} = \frac{1}{\sqrt{5}} \overset{\mathbf{K}}{\begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}} * \overset{\mathbf{L}}{\begin{bmatrix} 10 & 0 \\ 0 & 0 \end{bmatrix}} * \frac{1}{\sqrt{5}} \overset{\mathbf{K}^T}{\begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}}$$

$$\mathbf{R} = \overset{\mathbf{A}}{\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}} * \overset{\mathbf{A}^T}{\begin{bmatrix} 2 & 2 \\ 1 & 1 \end{bmatrix}} = \begin{bmatrix} 5 & 5 \\ 5 & 5 \end{bmatrix}$$

$$\mathbf{R} = \frac{1}{\sqrt{2}} \overset{\mathbf{K}}{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}} * \overset{\mathbf{L}}{\begin{bmatrix} 10 & 0 \\ 0 & 0 \end{bmatrix}} * \frac{1}{\sqrt{2}} \overset{\mathbf{K}^T}{\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}$$

Przykład rozkładu SVD

$$\mathbf{T} = \begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix}$$

$$\mathbf{U} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}$$

$$\mathbf{V} = \frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}$$

$$\underbrace{\frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{\mathbf{U}} * \underbrace{\begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix}}_{\mathbf{T}} * \underbrace{\frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}}_{\mathbf{V}^T} = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} = \mathbf{A}$$

Przykład rozkładu SVD

- Niesymetryczna macierz prostokątna 1x2

$$\mathbf{A} = \begin{bmatrix} 1 & 2 \end{bmatrix}$$

$$\mathbf{A} = \underbrace{\begin{bmatrix} 1 \end{bmatrix}}_{\mathbf{U}} * \underbrace{\begin{bmatrix} \sqrt{5} & 0 \end{bmatrix}}_{\mathbf{T}} * \underbrace{\frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}}_{\mathbf{V}^T}$$

Przykład rozkładu SVD

$$\mathbf{C} = \overset{\mathbf{A}^T}{\begin{bmatrix} 1 \\ 2 \end{bmatrix}} * \overset{\mathbf{A}}{\begin{bmatrix} 1 & 2 \end{bmatrix}} = \begin{bmatrix} 1 & 2 \\ 2 & 4 \end{bmatrix}$$

$$\mathbf{C} = \overset{\mathbf{K}}{\frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}} * \overset{\mathbf{L}}{\begin{bmatrix} 5 & 0 \\ 0 & 0 \end{bmatrix}} * \overset{\mathbf{K}^T}{\frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}}$$

$$\mathbf{R} = \overset{\mathbf{A}}{\begin{bmatrix} 1 & 2 \end{bmatrix}} * \overset{\mathbf{A}^T}{\begin{bmatrix} 1 \\ 2 \end{bmatrix}} = \begin{bmatrix} 5 \end{bmatrix}$$

$$\mathbf{R} = \overset{\mathbf{K}}{\begin{bmatrix} 1 \end{bmatrix}} * \overset{\mathbf{L}}{\begin{bmatrix} 5 \end{bmatrix}} * \overset{\mathbf{K}^T}{\begin{bmatrix} 1 \end{bmatrix}}$$

Przykład rozkładu SVD

T= $\begin{bmatrix} \sqrt{5} & 0 \end{bmatrix}$

U= $\begin{bmatrix} 1 \end{bmatrix}$

V= $\frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}$

$$\underset{\mathbf{U}}{\begin{bmatrix} 1 \end{bmatrix}} * \underset{\mathbf{T}}{\begin{bmatrix} \sqrt{5} & 0 \end{bmatrix}} * \underset{\mathbf{V}^T}{\frac{1}{\sqrt{5}} \begin{bmatrix} 1 & 2 \\ 2 & -1 \end{bmatrix}} = \begin{bmatrix} 1 & 2 \end{bmatrix} = \mathbf{A}$$

Uogólnienie EVD: rozkład SVD

- Szczególnym przypadkiem rozkładu SVD jest rozkład EVD
  - dla macierzy symetrycznej zachodzi: $\mathbf{A}^T = \mathbf{A}$, z czego wynika, że:

    $\mathbf{C} = \mathbf{A}^T\mathbf{A} = \mathbf{A}\mathbf{A} = \mathbf{A}^2$

    $\mathbf{R} = \mathbf{A}\mathbf{A}^T = \mathbf{A}\mathbf{A} = \mathbf{A}^2$
    - czyli $\mathbf{C} = \mathbf{R}$
  - obliczając $\mathbf{A}^2$ z wykorzystaniem rozkładu $\mathbf{A} = \mathbf{KLK}^T$ mamy:

    $\mathbf{A}^2 = \mathbf{KLK}^T\mathbf{KLK}^T = \mathbf{KLLK}^T = \mathbf{KL}^2\mathbf{K}^T$
  - wniosek:
    - wartościami własnymi macierzy $\mathbf{A}^2$ są kwadraty wartości własnych macierzy $\mathbf{A}$, a więc wartości osobliwymi symetrycznej macierzy $\mathbf{A}$ są wartości własne (ze znakiem plus) macierzy $\mathbf{A}$
    - wektorami osobliwymi symetrycznej macierzy $\mathbf{A}$ są jej wektory własne

Uogólnienie EVD: rozkład SVD

- Podsumowanie:
  - po rozłożeniu symetrycznej macierzy **A** na $\mathbf{KLK}^T$ (rozkład EVD) oraz $\mathbf{UTV}^T$ (rozkład SVD) mamy:

    $\mathbf{A} = \mathbf{KLK}^T = \mathbf{UTV}^T$

    - jednocześnie zachodzi:

    $\mathbf{U} = \mathbf{K}$ oraz $\mathbf{V}^T = \mathbf{K}^T$

    - z czego wynika, że **U** = **V**
  - ale uwaga:
    - **K** (a tym samym **U** i **V**) są określone z dokładnością do znaku kolumn (czyli znaku zawartych w nich wektorów własnych)

Uogólnienie EVD: rozkład SVD

- Problem ze znakiem wektorów własnych
  - wektory własne macierzy są określone z dokładnością do znaku, tzn. jeżeli zachodzi: $\mathbf{A} = \mathbf{KLK}^T$, gdzie $\mathbf{K}$ jest macierzą ortogonalną (czyli $\mathbf{K}^T\mathbf{K} = \mathbf{I}$) a dodatkowo:
    - macierz $\mathbf{Q}$ różni się od macierzy $\mathbf{K}$ znakiem dowolnej liczby kolumn ($\mathbf{Q} = \mathbf{K} \cdot \mathrm{diag}(\pm 1, \pm 1, ..., \pm 1)$ – prawostronne mnożenie przez macierz diagonalną zawierającą –1 zmienia znak odpowiedniej kolumny)
  - to zachodzi także $\mathbf{A} = \mathbf{QLQ}^T$ oraz $\mathbf{Q}^T\mathbf{Q} = \mathbf{I}$
- Niejednoznaczność znaku kolumn macierzy $\mathbf{K}$ nie ma wpływu na $\mathbf{KLK}^T$ (rozkład EVD macierzy $\mathbf{C} = \mathbf{A}^T\mathbf{A}$), ma jednak wpływ na $\mathbf{UTV}^T$, gdzie za $\mathbf{V}$ przyjmuje się $\mathbf{K}$ (rozkład SVD macierzy $\mathbf{A}$), czyli:
  - jeżeli $\mathbf{Q} = \mathbf{K} \cdot \mathrm{diag}(\pm 1, \pm 1, ..., \pm 1)$ ($\mathbf{Q}$ różni się od $\mathbf{K}$ znakami kolumn), to:

    $\mathbf{KLK}^T = \mathbf{QLQ}^T$, ale $\mathbf{UTK}^T \neq \mathbf{UTQ}^T$ (macierze $\mathbf{K}$ i $\mathbf{Q}$ podstawiono pod $\mathbf{V}$ w wyrażeniu $\mathbf{UTV}^T$)
- Sytuacja wygląda analogicznie, gdy za $\mathbf{U}$ przyjmuje się macierz $\mathbf{K}$ wynikającą z rozkładu EVD macierzy $\mathbf{R} = \mathbf{AA}^T$
  - jeżeli $\mathbf{Q} = \mathbf{K} \cdot \mathrm{diag}(\pm 1, \pm 1, ..., \pm 1)$ to $\mathbf{KLK}^T = \mathbf{QLQ}^T$, ale $\mathbf{KTV}^T \neq \mathbf{QTV}^T$

Przykład rozkładu SVD

- Przykład: macierz **A** oraz odpowiadająca jej **C** = $\mathbf{A}^T\mathbf{A}$

$$\mathbf{A}=\begin{bmatrix}2 & 1\\ 2 & 1\end{bmatrix} \qquad \mathbf{C}=\underbrace{\begin{bmatrix}2 & 2\\ 1 & 1\end{bmatrix}}_{\mathbf{A}^T} * \underbrace{\begin{bmatrix}2 & 1\\ 2 & 1\end{bmatrix}}_{\mathbf{A}} = \begin{bmatrix}8 & 4\\ 4 & 2\end{bmatrix}$$

  - rozkład macierzy **C** na $\mathbf{KLK}^T$ jest niejednoznaczny:
    - zmiana znaku dowolnej kolumny w macierzy **K** (a tym samym: wiersza macierzy $\mathbf{K}^T$) nie zmienia wartości wyrażenia $\mathbf{KLK}^T$

$$\underbrace{\frac{1}{\sqrt{5}}\begin{bmatrix}2 & 1\\ 1 & -2\end{bmatrix}}_{\mathbf{K}} * \underbrace{\begin{bmatrix}10 & 0\\ 0 & 0\end{bmatrix}}_{\mathbf{L}} * \underbrace{\frac{1}{\sqrt{5}}\begin{bmatrix}2 & 1\\ 1 & -2\end{bmatrix}}_{\mathbf{K}^T} = \begin{bmatrix}8 & 4\\ 4 & 2\end{bmatrix} = \mathbf{C}$$

$$\underbrace{\frac{1}{\sqrt{5}}\begin{bmatrix}-2 & 1\\ -1 & -2\end{bmatrix}}_{\mathbf{Q}} * \underbrace{\begin{bmatrix}10 & 0\\ 0 & 0\end{bmatrix}}_{\mathbf{L}} * \underbrace{\frac{1}{\sqrt{5}}\begin{bmatrix}-2 & -1\\ 1 & -2\end{bmatrix}}_{\mathbf{Q}^T} = \begin{bmatrix}8 & 4\\ 4 & 2\end{bmatrix} = \mathbf{C}$$

  - itd.

Przykład rozkładu SVD

- Przykład: c.d.
  - niejednoznaczność znaku kolumn macierzy **K** nie ma wpływu na wynik wyrażenia **KLK**^{T} (rozkład EVD macierzy **C** = **A**^{T}**A**), ma jednak wpływ na wynik wyrażenia **UTV**^{T}, gdzie za **V** przyjmuje się **K** (rozkład SVD macierzy **A**)

$$\underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}}_{\mathbf{U}} * \underbrace{\begin{bmatrix}\sqrt{10} & 0\\ 0 & 0\end{bmatrix}}_{\mathbf{T}} * \underbrace{\frac{1}{\sqrt{5}}\begin{bmatrix}2 & 1\\ 1 & -2\end{bmatrix}}_{\mathbf{V}^T=\mathbf{K}^T} = \begin{bmatrix}2 & 1\\ 2 & 1\end{bmatrix} = \mathbf{A}$$

$$\underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}}_{\mathbf{U}} * \underbrace{\begin{bmatrix}\sqrt{10} & 0\\ 0 & 0\end{bmatrix}}_{\mathbf{T}} * \underbrace{\frac{1}{\sqrt{5}}\begin{bmatrix}-2 & -1\\ 1 & -2\end{bmatrix}}_{\mathbf{V}^T=\mathbf{Q}^T} = \begin{bmatrix}-2 & 1\\ -2 & 1\end{bmatrix} \neq \mathbf{A}$$

Uogólnienie EVD: rozkład SVD

- Problem niejednoznaczności znaku kolumn można rozwiązać wykorzystując dwie postaci zależności pomiędzy wektorami własnymi macierzy $\mathbf{A}^T\mathbf{A}$ oraz $\mathbf{A}\mathbf{A}^T$

$\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{V}^T$ // (prawostr.) * $\mathbf{V}$

$\mathbf{AV} = \mathbf{U}\mathbf{T}\mathbf{V}^T\mathbf{V}$

$\mathbf{AV} = \mathbf{UTI}$

$\mathbf{AV} = \mathbf{UT}$

$\mathbf{AV} = \mathbf{UT}$ // (prawostr.) * $(\mathbf{T}_{\square})^{\sim1}$

$\mathbf{AV}(\mathbf{T}_{\square})^{\sim1} = \mathbf{UT}(\mathbf{T}_{\square})^{\sim1}$

$\mathbf{A} = \mathbf{U}\mathbf{T}\mathbf{V}^T$ // (lewostr.) * $\mathbf{U}^T$

$\mathbf{U}^T\mathbf{A} = \mathbf{U}^T\mathbf{U}\mathbf{T}\mathbf{V}^T$

$\mathbf{U}^T\mathbf{A} = \mathbf{I}\mathbf{T}\mathbf{V}^T$

$\mathbf{U}^T\mathbf{A} = \mathbf{T}\mathbf{V}^T$

$\mathbf{U}^T\mathbf{A} = \mathbf{T}\mathbf{V}^T$ // (lewostr.) * $(\mathbf{T}_{\square})^{\sim1}$

$(\mathbf{T}_{\square})^{\sim1}\mathbf{U}^T\mathbf{A} = (\mathbf{T}_{\square})^{\sim1}\mathbf{T}\mathbf{V}^T$

  - dla osobliwych **T** powyższy „wywód” musi być tutaj zatrzymany, co oznacza, że utworzyliśmy jedynie macierze $\mathbf{UT}(\mathbf{T}_{\square})^{\sim1}$ (zamiast **U**) względnie $(\mathbf{T}_{\square})^{\sim1}\mathbf{T}\mathbf{V}^T$ (zamiast $\mathbf{V}^T$)
    - macierze te dzięki czynnikowi $\mathbf{T}(\mathbf{T}_{\square})^{\sim1}$ będą miały zerowe kolumny względnie wiersze, będą więc wprawdzie mogły być użyte w konstruowanym rozkładzie, ale nie będą macierzami ortogonalnymi
  - dla nieosobliwych **T** zachodzi $\mathbf{T}(\mathbf{T}_{\square})^{\sim1} = \mathbf{I}$, co oznacza, że możliwe jest wykonanie redukcji po prawych stronach powyższych równań, co prowadzi do ostatecznego powstania poszukiwanych macierzy **U** i $\mathbf{V}^T$:

$\mathbf{AV}(\mathbf{T}_{\square})^{\sim1} = \mathbf{U}$

$(\mathbf{T}_{\square})^{\sim1}\mathbf{U}^T\mathbf{A} = \mathbf{V}^T$

Uogólnienie EVD: rozkład SVD

- Uwaga: zapis $(\mathbf{T}_{\square})^{\sim 1}$ oznacza złożenie dwóch operacji:
  - $\mathbf{X}_{\square}$ oznacza pewną nieformalną operację służącą do przekształcenia prostokątnej macierzy **X** w macierz kwadratową, która polega na:
    - usunięciu ostatnich wierszy (jeżeli liczba ta przekracza liczbę kolumn)
    - usunięciu ostatnich kolumn (jeżeli liczba ta przekracza liczbę wierszy)
  - z kolei zapis $\mathbf{X}^{\sim 1}$ oznacza pseudo-odwrotność macierzy **X**
    - jeżeli $\mathbf{X} = \mathrm{diag}([x_{11}, x_{22}, ..., x_{nn}])$ to $\mathbf{X}^{\sim 1} = \mathrm{diag}([y_{11}, y_{22}, ..., y_{nn}])$, przy czym
      - $y_{ii} = 1/x_{ii}$ dla $x_{ii} \neq 0$
      - $y_{ii} = 0$ dla $x_{ii} = 0$

Uogólnienie EVD: rozkład SVD

- Wniosek: **U** i **V** muszą być tak dobrane, aby zachodziło: **AV** = **UT** oraz $\mathbf{U}^T\mathbf{A}$ = $\mathbf{T}\mathbf{V}^T$, można więc:
  - ustalić **U** jako macierz wektorów własnych dla $\mathbf{A}\mathbf{A}^T$ oraz **V** jako macierz wektorów własnych dla $\mathbf{A}^T\mathbf{A}$, a następnie zmieniać znaki kolumn macierzy **U** i **V** aż do osiągnięcia $\mathbf{U}\mathbf{T}\mathbf{V}^T$ = **A**

Uogólnienie EVD: rozkład SVD

- Lepszy (choć bardziej złożony) pomysł
  - gdy **A** ma rozmiary mxn, przy czym m > n
    - najpierw utworzyć macierz **V** a potem wykorzystać ją do utworzenia macierzy **U** wykonując operację: $\mathbf{AV}(\mathbf{T}_{\neq 0})^{-1}$
      - jeżeli wszystkie diagonalne elementy macierzy **T** są różne od 0, to operacja ta generuje n początkowych kolumn macierzy **U**
      - jeżeli r < n diagonalnych elementów macierzy **T** wynosi 0, to operacja ta generuje r początkowych kolumn macierzy **U** i n–r wektorów zerowych (kolumny zerowe nie mogą być użyte jako kolumny macierzy **U**)
    - macierz **U** należy utworzyć wykonując procedurę ortogonalizacji Grama-Schmidt'a na macierzy wektorów $[\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_p, \mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_r]$, gdzie:
      - $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_p$ są niezerowymi kolumnami macierzy $\mathbf{AV}(\mathbf{T}_{\neq 0})^{-1}$
        - » uwaga: $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_p\}$ jest zbiorem wektorów niezależnych
      - $\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_r$ jest zbiorem dowolnych wektorów takich, że $\{\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_p, \mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_r\}$ jest zbiorem wektorów niezależnych
        - » wektory $\mathbf{w}_1, \mathbf{w}_2, ..., \mathbf{w}_r$ są więc niezależne od siebie oraz od wektorów $\mathbf{u}_1, \mathbf{u}_2, ..., \mathbf{u}_p$
      - dowolnymi wektorami, przy czym r = m–p
  - gdy **A** ma rozmiary mxn, przy czym m < n
    - tworzy się najpierw macierz **U**, a na jej podstawie macierz **V**

Przykład rozkładu SVD

- Symetryczna macierz kwadratowa 2x2

$$\mathbf{A}=\begin{bmatrix}2 & 1\\ 2 & 1\end{bmatrix}\quad \mathbf{C}=\begin{bmatrix}8 & 4\\ 4 & 2\end{bmatrix}\quad \mathbf{R}=\begin{bmatrix}5 & 5\\ 5 & 5\end{bmatrix}$$

$$\mathbf{C}=\underbrace{\frac{1}{\sqrt{5}}\begin{bmatrix}2 & 1\\ 1 & -2\end{bmatrix}}_{\mathbf{K}} * \underbrace{\begin{bmatrix}10 & 0\\ 0 & 0\end{bmatrix}}_{\mathbf{L}} * \underbrace{\frac{1}{\sqrt{5}}\begin{bmatrix}2 & 1\\ 1 & -2\end{bmatrix}}_{\mathbf{K}^T}$$

$$\mathbf{R}=\underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}}_{\mathbf{K}} * \underbrace{\begin{bmatrix}10 & 0\\ 0 & 0\end{bmatrix}}_{\mathbf{L}} * \underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}}_{\mathbf{K}^T}$$

$$\mathbf{A}=\underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix}1 & 1\\ 1 & -1\end{bmatrix}}_{\mathbf{U}} * \underbrace{\begin{bmatrix}\sqrt{10} & 0\\ 0 & 0\end{bmatrix}}_{\mathbf{T}} * \underbrace{\frac{1}{\sqrt{5}}\begin{bmatrix}2 & 1\\ 1 & -2\end{bmatrix}}_{\mathbf{V}^T}$$

Przykład rozkładu SVD

- Symetryczna macierz kwadratowa 2x2 c.d.
  - założmy, że obliczono najpierw **C**=**A**^{T}**A** oraz jej rozkład **KLK**^{T}

$$
\mathbf{C} = \underset{\mathbf{K}}{\frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}} * \underset{\mathbf{L}}{\begin{bmatrix} 10 & 0 \\ 0 & 0 \end{bmatrix}} * \underset{\mathbf{K}^T}{\frac{1}{\sqrt{5}} \begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}}
$$

    - obliczona w ten sposób macierz **K** będzie pełniła rolę macierzy **V**
    - natomiast obliczona w ten sposób macierz **L** posłuży do utworzenia macierzy **T** (wynikowa macierz **T** byłaby identyczna, gdyby do jej utworzenia zastosować macierz **L** powstałą podczas rozkładu macierzy **R** = **AA**^{T})

Przykład rozkładu SVD

- Symetryczna macierz kwadratowa 2x2 c.d.
  - obliczamy **AV**

$$\mathbf{AV} = \underset{\mathbf{A}}{\begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix}} * \frac{1}{\sqrt{5}} \underset{\mathbf{V}}{\begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}} = \frac{1}{\sqrt{5}} \begin{bmatrix} 5 & 0 \\ 5 & 0 \end{bmatrix}$$

  - tworzymy macierz **T**

$$\mathbf{T} = \begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix}$$

  - i macierz $\mathbf{T}_{\square}$ = **T** (ponieważ **T** jest macierzą kwadratową)

$$\mathbf{T}_{\square} = \begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix}$$

Przykład rozkładu SVD

- Symetryczna macierz kwadratowa 2x2 c.d.
  - obliczamy $(\mathbf{T}_{\square})^{\sim 1}$

$$\mathbf{T} = \frac{1}{\sqrt{10}} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}$$

  - obliczamy $\mathbf{AV}(\mathbf{T}_{\neq 0})^{\sim 1}$

$$\mathbf{AV}(\mathbf{T}_{\neq 0})^{\sim 1} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 0 \\ 1 & 0 \end{bmatrix}$$

  - macierz $\mathbf{AV}(\mathbf{T}_{\neq 0})^{\sim 1}$ zawiera jedną niezerową kolumnę
    - jest to pierwsza kolumna tworzonej macierzy **U**
    - pozostałe kolumny tworzonej macierzy **U** mogą być dowolnymi wektorami, które „uzupełnią” **U** do macierzy ortogonalnej (tzn. po dodaniu tych kolumn **U** musi spełniać $\mathbf{U}^T\mathbf{U} = \mathbf{I}$)

Przykład rozkładu SVD

- Symetryczna macierz kwadratowa 2x2 c.d.
  - unormowanymi wektorami ortogonalnymi do wektora:

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

  - są wektory:

$$\mathbf{v} = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 \\ -1 \end{bmatrix} \quad \text{oraz} \quad \mathbf{w} = -\mathbf{v} = \frac{1}{\sqrt{2}} \begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

  - ostatecznie macierz **U** może przyjąć jedną z postaci:

$$\mathbf{U}_1 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad \text{lub} \quad \mathbf{U}_2 = \frac{1}{\sqrt{2}} \begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$$

Przykład rozkładu SVD

- Symetryczna macierz kwadratowa 2x2 c.d.
  - dla obu postaci (tzn. $\mathbf{U}_1$, $\mathbf{U}_2$) tak ustalonej macierzy $\mathbf{U}$ zachodzi:

$$\underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix}}_{\mathbf{U}_1} * \underbrace{\begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix}}_{\mathbf{T}} * \underbrace{\frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}}_{\mathbf{V}^T} = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} = \mathbf{A}$$

$$\underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}}_{\mathbf{U}_2} * \underbrace{\begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix}}_{\mathbf{T}} * \underbrace{\frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}}_{\mathbf{V}^T} = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} = \mathbf{A}$$

    - uwaga: że względu na zerowy element przekątnej macierzy **T** druga kolumna **U** jest określona na tyle niejednoznacznie, że może przyjąć dowolne (niekoniecznie unormowane) wartości

$$\underbrace{\frac{1}{\sqrt{2}}\begin{bmatrix} 1 & 250 \\ 1 & -97 \end{bmatrix}}_{\mathbf{U}_2} * \underbrace{\begin{bmatrix} \sqrt{10} & 0 \\ 0 & 0 \end{bmatrix}}_{\mathbf{T}} * \underbrace{\frac{1}{\sqrt{5}}\begin{bmatrix} 2 & 1 \\ 1 & -2 \end{bmatrix}}_{\mathbf{V}^T} = \begin{bmatrix} 2 & 1 \\ 2 & 1 \end{bmatrix} = \mathbf{A}$$

Przykład rozkładu SVD

- Symetryczna macierz kwadratowa 2x2 c.d.
  - niejednoznaczność drugiej kolumny macierzy **U**
    - wynika ona z faktu, że macierz **U** mnożona jest przez macierz **T**, w której drugi element głównej przekątnej wynosi zero
    - element ten „zeruje” więc drugą warstwę iloczynu (która powstaje z drugiej kolumny macierzy **U** i drugiego wiersza macierzy $\mathbf{V}^T$)
    - oznacza to jednocześnie, że (z tego samego powodu) drugi wiersz macierzy $\mathbf{V}^T$ jest także określony niejednoznacznie
    - wniosek: druga kolumna macierzy **U** i drugi wiersz macierzy $\mathbf{V}^T$ nie są ze sobą związane i mogą przyjmować zupełnie dowolne wartości
    - w praktyce oczywiście dobiera je jednak się tak, żeby:
      - pierwsza kolumna macierzy **U** była unormowana i ortogonalna do pozostałych kolumn tej macierzy ($\mathbf{U}^T\mathbf{U} = \mathbf{I}$)
      - pierwszy wiersz macierzy $\mathbf{V}^T$ był unormowany i ortogonalny do pozostałych wierszy tej macierzy ($\mathbf{V}^T\mathbf{V} = \mathbf{I}$)

Przykład rozkładu SVD

- Symetryczna macierz kwadratowa 2x2 c.d.
  - niejednoznaczność powyższa nie dotyczy całej macierzy **U**, a jedynie drugiej kolumny
  - w odróżnieniu od niej pierwsza kolumny macierzy **U** jest wyznaczona jednoznacznie
    - pierwsza kolumna macierzy **U** jest związana z pierwszym wierszem macierzy $\mathbf{V}^T$ zależnością **AV** = **UT**, co oznacza, że zarówno pierwsza kolumna macierzy **U** jak i pierwszy wiersz macierzy $\mathbf{V}^T$ muszą być tak dobrane, aby zachodziło jednocześnie:
      - pierwsza kolumna macierzy **U** musi być unormowana i ortogonalna do pozostałych kolumn tej macierzy ($\mathbf{U}^T\mathbf{U} = \mathbf{I}$)
      - pierwszy wiersz macierzy $\mathbf{V}^T$ musi być unormowany i ortogonalny do pozostałych wierszy tej macierzy ($\mathbf{V}^T\mathbf{V} = \mathbf{I}$)
      - pomiędzy znakami pierwszej kolumny macierzy **U** i pierwszego wiersza macierzy $\mathbf{V}^T$ musi zachodzić zgodność (**AV** = **UT**)

…

Stratna kompresja obrazu

- Dzięki rozkładowi **A** = **UTV**^{T} możliwe jest tworzenie różnych przybliżeń dowolnej macierzy **A** (elementami skalującymi w tym iloczynie są tzw. wartości osobliwe macierzy **A**, umieszczone na przekątnej macierzy **T**)
- Niech rozmiarem macierzy **A** będzie mxn, wtedy odpowiednie rozmiary czynników iloczynu przedstawiają się następująco:
  - **U**: mxm
  - **T**: mxn
  - **V**: nxn
- Do zapamiętania
  - macierzy **A** potrzeba m·n wartości
  - macierzy **U**, **T** i **V**^{T} potrzeba m·m + min(m,n) + n·n wartości
  - przybliżenia macierzy **A** za pomocą p składowych (p < min(m,n)), czyli:
    - p kolumn macierzy **U**,
    - p wartości macierzy **T**
    - p wierszy macierzy **V**

    potrzeba p·(m + n + 1) wartości

Macierz $\mathbf{UTV}^T$ jako obraz

- Ponieważ obraz (dwuwymiarowy) można traktować jako pewną macierz, obrazy mogą być poddawane operacjom macierzowym
- Takie traktowanie danych obrazowych ma wyraźną przewagę nad postępowaniem polegającym na klasycznym stosowaniu PCA
  - pomija fazę tworzenia cech opisujących obraz (np. zmiennych RGB)
  - w naturalny sposób traktuje obraz jako strukturę dwuwymiarową
    - uwaga: obrazu nie trzeba traktować jako struktury dwuwymiarowej, szczególnie wtedy, gdy dysponuje się dobrymi cechami wyższych rzędów, tak jest jednak bardzo rzadko, dlatego traktowanie obrazu jako struktury dwuwymiarowej jest dobrym kompromisem
- Bardzo szczególne obrazy (kwadratowe i symetryczne), przedstawione w postaci macierzy **A**, mogą być przetwarzane dzięki rozkładowi $\mathbf{A} = \mathbf{KLK}^T$ (metoda EVD)
- W ogólności jednak do przetwarzania obrazów (niekoniecznie kwadratowych i symetrycznych), przedstawionych w postaci macierzy **B**, stosuje się (ogólniejszy) rozkład $\mathbf{B} = \mathbf{UTV}^T$ (metoda SVD)

Przykłady rozkładu macierzy-obrazu

- Obraz oryginalny #1 -- definicja:
  A1 = $[a_{ij}]$, i = 1..100, j = 1..100, gdzie:
  $a_{ij} = \sin(\pi/50 \cdot i)\sin(\pi/50 \cdot j)$

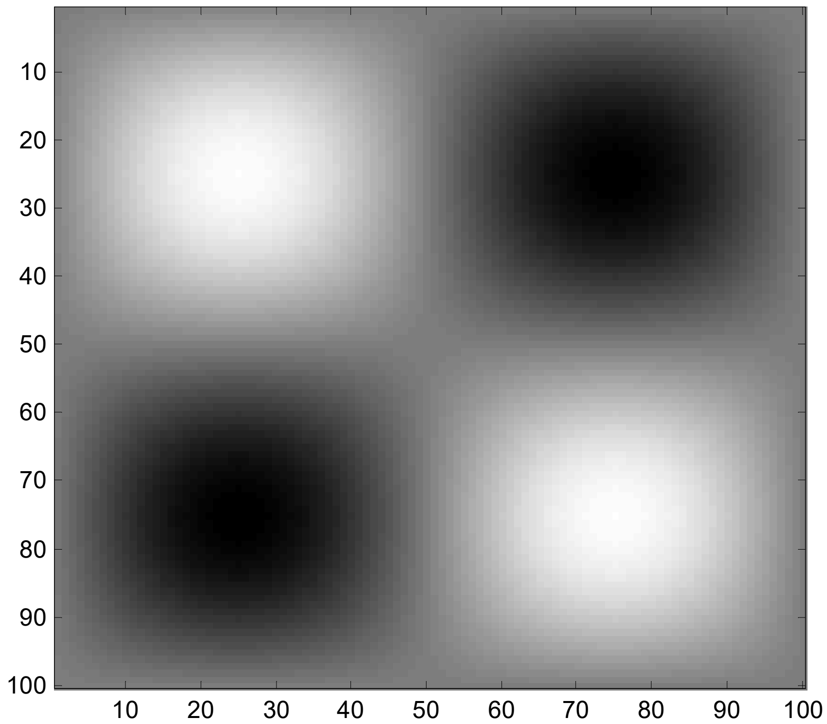
- Obraz oryginalny #2 -- definicja:
  A2 = **A1** + zakłócenia

- Obraz oryginalny #2 -- definicja:
  A3 = **A1** + **F**, gdzie **F** jest pewną macierzą (przedstawiającą rodzaj drobnej „kratki")

(rozmiary we float'ach [4-bajtowych reprezentacjach liczb rzeczywistych])

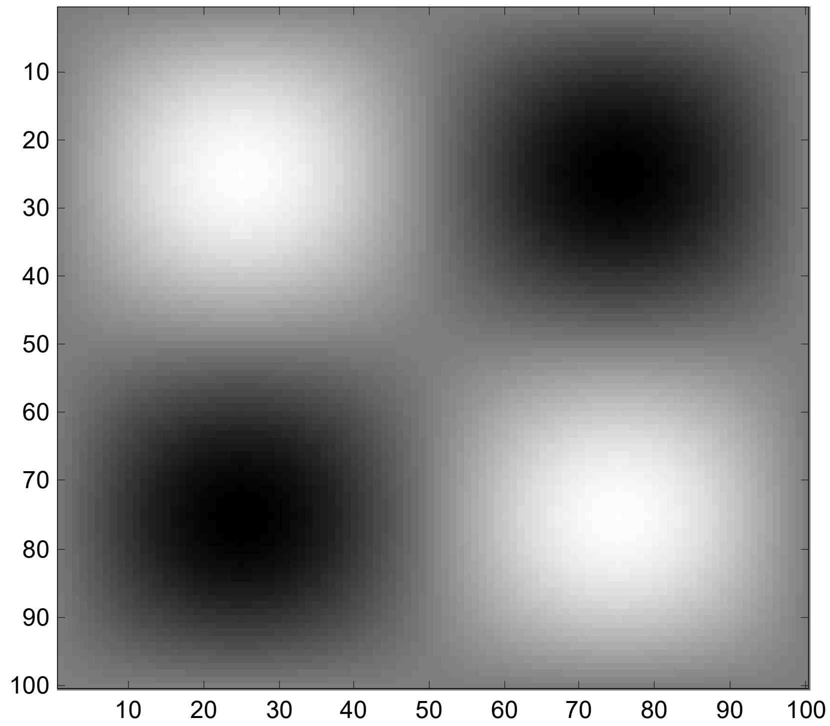
Rozkład obrazu #1, metoda SVD

- Obraz oryginalny (100x100, 100*100 = 10000f)



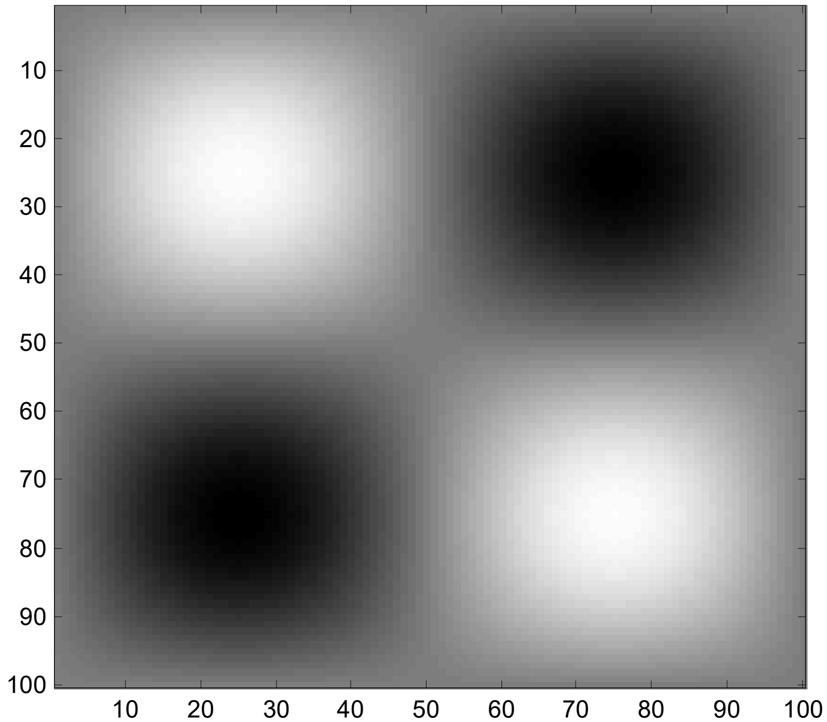
Rozkład obrazu #1, metoda SVD

- 100 warstw (20100f, 201.00% oryginału)



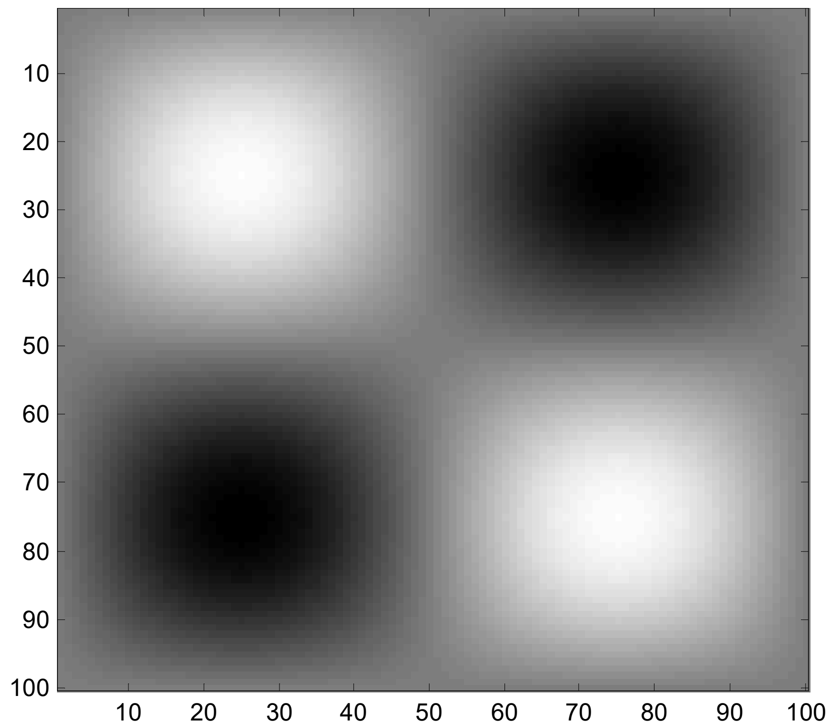
Rozkład obrazu #1, metoda SVD

- 10 warstw (2010f, 20.10% oryginału)



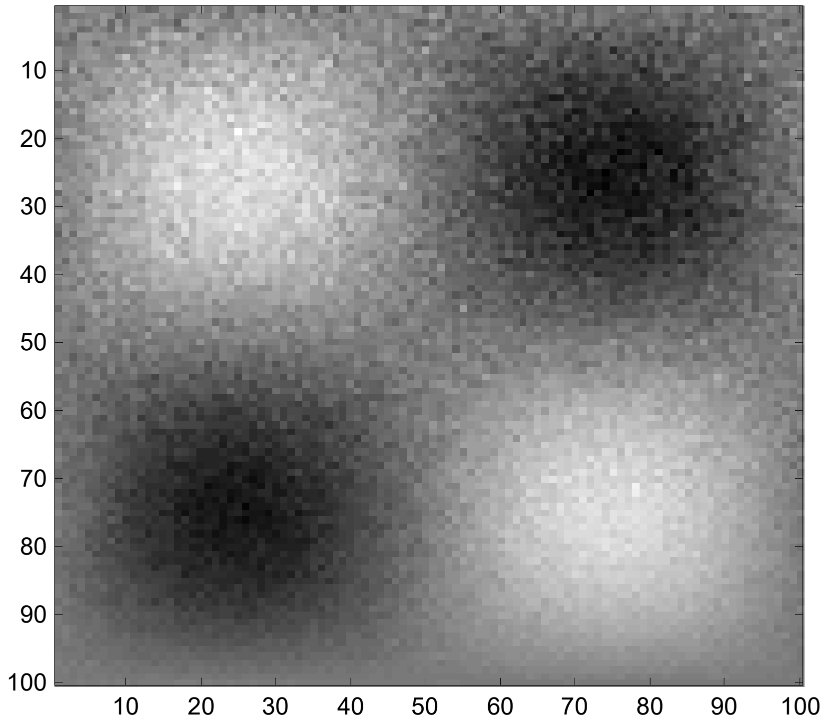
Rozkład obrazu #1, metoda SVD

- 1 warstwa (201f, 2.01% oryginału)



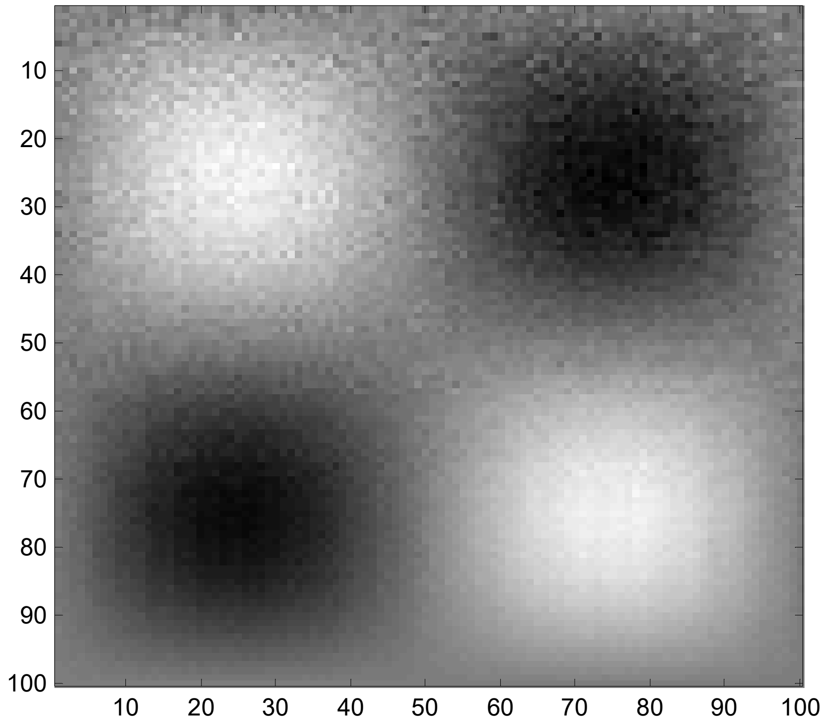
Rozkład obrazu #2, metoda SVD

- Obraz oryginalny (100x100, 100*100 = 10000f)



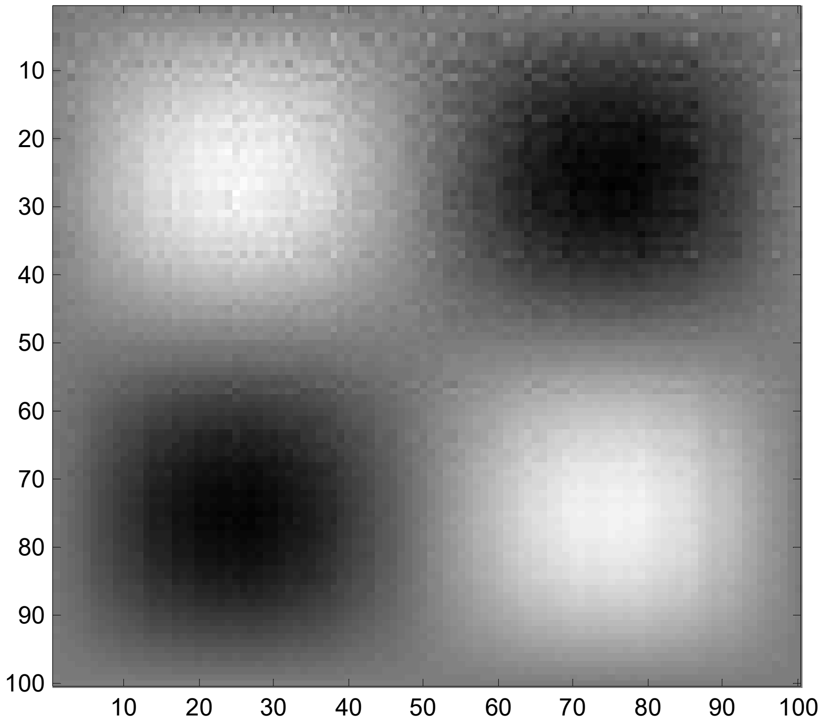
Rozkład obrazu #2, metoda SVD

- 10 warstw (2010f, 20.10% oryginału)



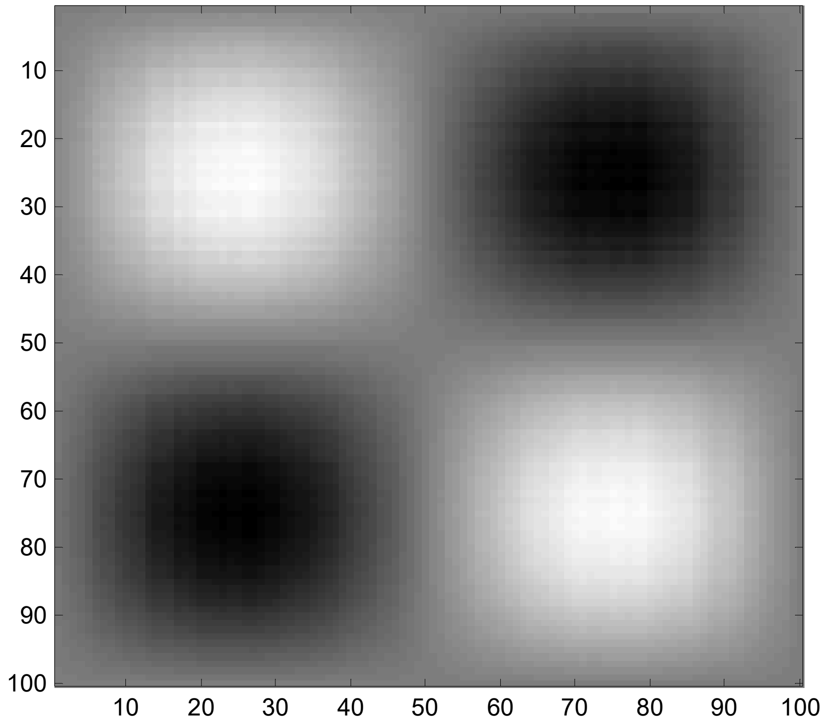
Rozkład obrazu #2, metoda SVD

- 3 warstwy (603f, 6.03% oryginału)



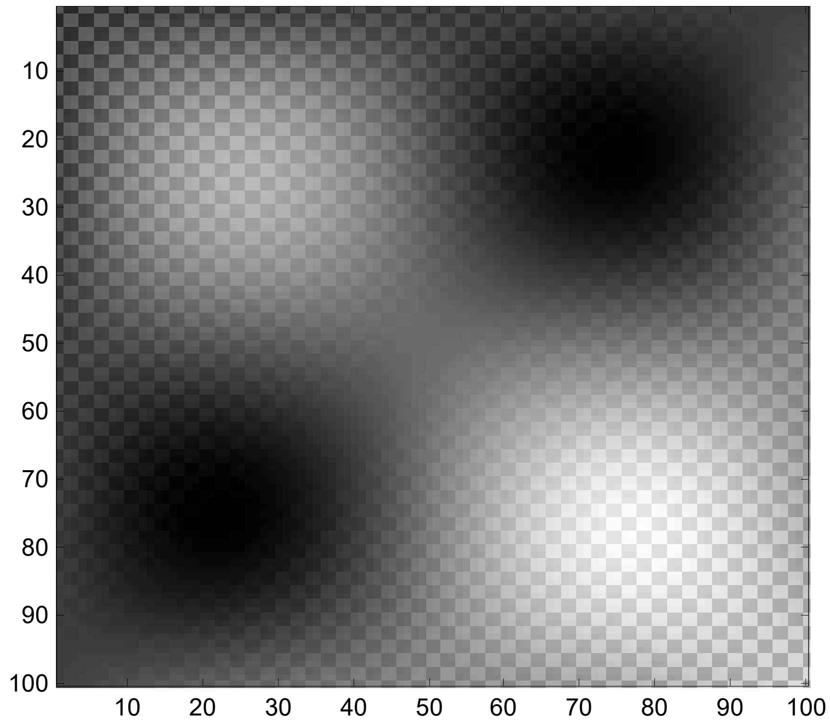
Rozkład obrazu #2, metoda SVD

- 1 warstwa (201f, 2.01% oryginału)



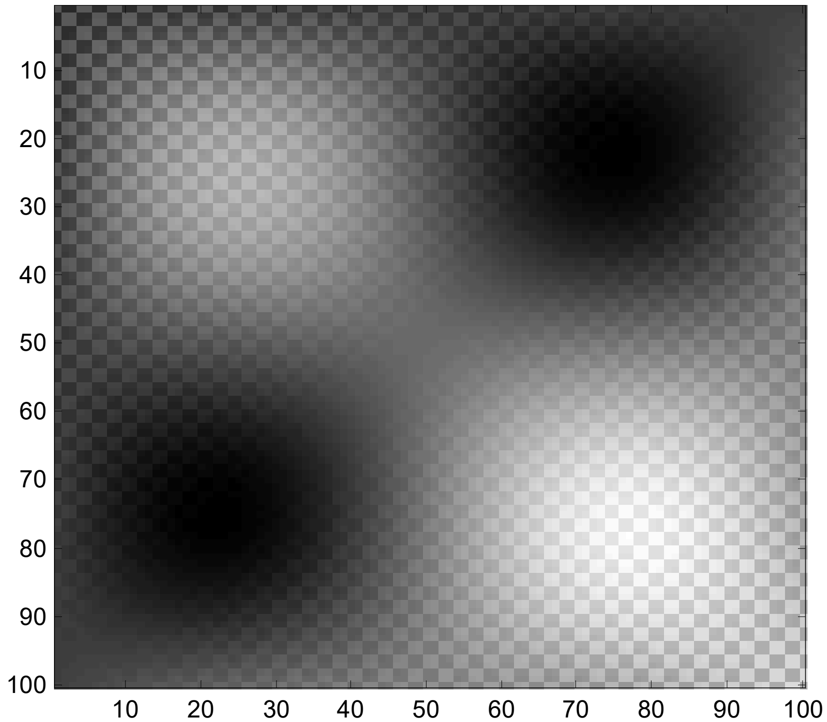
Rozkład obrazu #3, metoda SVD

- Obraz oryginalny (100x100, 100*100 = 10000f)



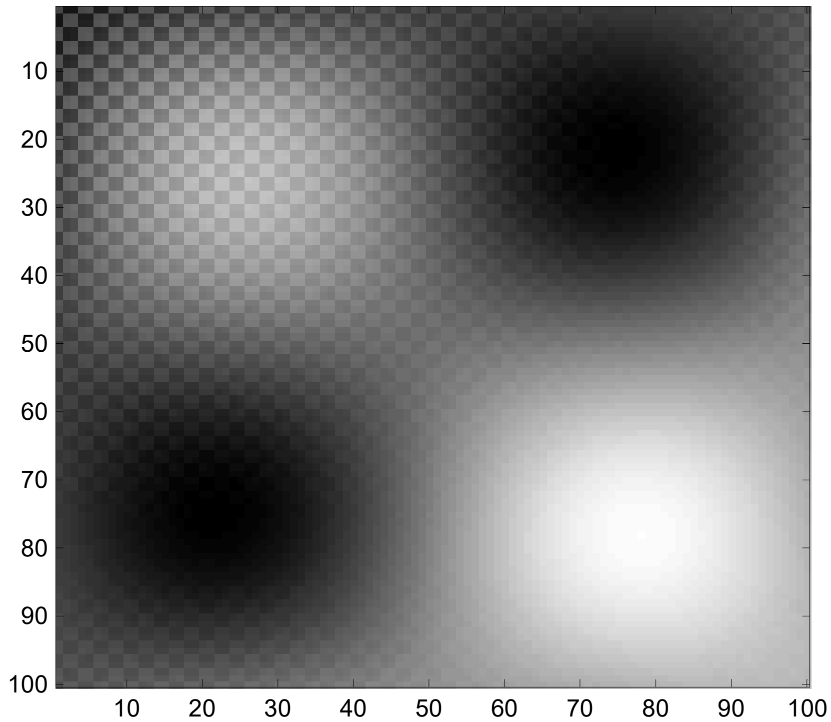
Rozkład obrazu #3, metoda SVD

- 6 warstw (1206f, 12.06% oryginału)



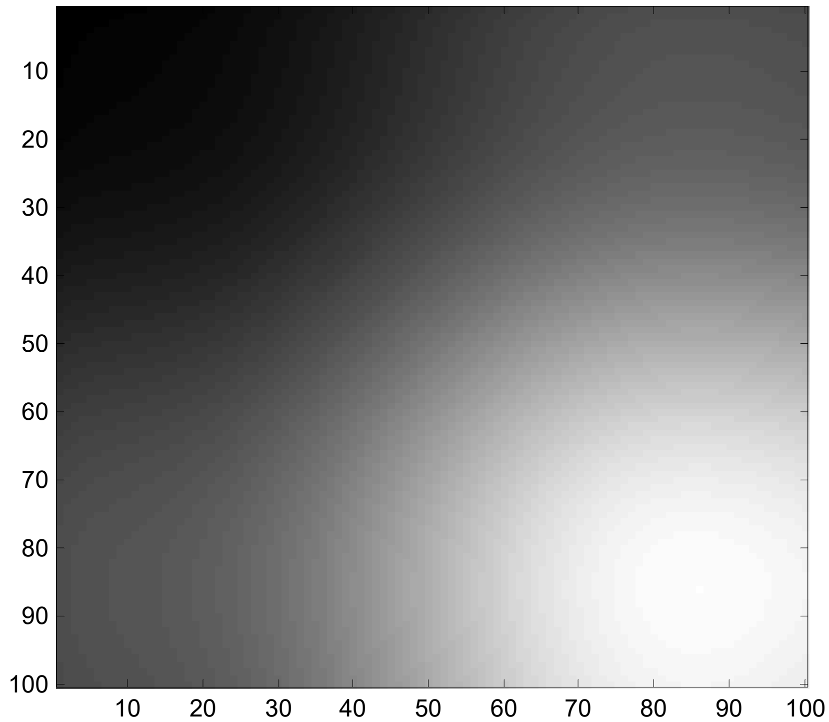
Rozkład obrazu #3, metoda SVD

- 3 warstwy (603f, 6.03% oryginału)



Rozkład obrazu #3, metoda SVD

- 1 warstwa (201f, 2.01% oryginału)



…

EVD a SVD

- Macierze określone
  - nieujemnie
  - dodatnio
  - niedodatnio
  - ujemnie
- Równoważność EVD i SVD
  - jeżeli **A** jest macierzą nieujemnie określoną, to jej rozkłady EVD i SVD są tożsame

…