...

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

...

Dygresja

- Pytanie: Skąd wartości własne λ_i macierzy **A** wiedzą, że mają spełniać tak wiele różnych właściwości? ☺
  - np. skąd wynikają zależności
    - trace(**A**) = $\sum\lambda_i$
    - det(**A**) = $\prod\lambda_i$
- Odpowiedź:
  - wartości własne są pierwiastkami wielomianu, a pierwiastki wielomianów „z automatu” spełniają wiele właściwości
    - w szczególności: pierwiastki wielomianu stopnia n-tego spełniają n zależności wyrażanych przez tzw. wzory Viete’a

Dygresja

- Pytanie: Skąd wartości własne λ_i macierzy **A** wiedzą, że mają spełniać tak wiele różnych właściwości? ☺
  - np. skąd wynikają zależności
    - trace(**A**) = $\sum\lambda_i$
    - det(**A**) = $\prod\lambda_i$
- Odpowiedź:
  - wielomian stopnia n
    - jest wyrażeniem postaci
      $ax^n + bx^{n-1} + ... + yx^1 + zx^0$ (czyli: $ax^n + bx^{n-1} + ... + yx + z$)
      - gdzie $a \neq 0$
    - i może być przedstawiony w postaci
      $a(x^n + b'x^{n-1} + ... + y'x^1 + z'x^0)$
      - gdzie $b' = b/a$, $c' = c/a$, ..., $z' = z/a$

Dygresja

- Pytanie: Skąd wartości własne λ_i macierzy **A** wiedzą, że mają spełniać tak wiele różnych właściwości? ☺
  – np. skąd wynikają zależności
    - trace(**A**) = $\sum\lambda_i$
    - det(**A**) = $\prod\lambda_i$
- Odpowiedź:
  – wielomian stopnia n
    - po oznaczeniu jego n (niekoniecznie różnych od siebie) pierwiastków* przez x_1, x_2, ..., x_n, ten sam wielomian może być przedstawiony w postaci $a((x–x_1)(x–x_2)\ldots(x–x_n))$

* liczba tych pierwiastków: patrz tzw. zasadnicze twierdzenie algebry

Dygresja

- Pytanie: Skąd wartości własne λ_i macierzy **A** wiedzą, że mają spełniać tak wiele różnych właściwości? ☺
  - np. skąd wynikają zależności
    - trace(**A**) = $\sum\lambda_i$
    - det(**A**) = $\prod\lambda_i$
- Odpowiedź:
  - wielomian stopnia n
    - $(x-x_1)(x-x_2)\ldots(x-x_n)$ =
      - wymnażamy
    - $(x-x_1)(x-x_2)\ldots(x-x_n)$ = „suma 2^n nieuporządkowanych składników”
      - porządkujemy
    - $(x-x_1)(x-x_2)\ldots(x-x_n)$ = „suma n+1 uporządkowanych składników”

Dygresja

- Pytanie: Skąd wartości własne λ_i macierzy **A** wiedzą, że mają spełniać tak wiele różnych właściwości? ☺
  - np. skąd wynikają zależności
    - trace(**A**) = $\sum\lambda_i$
    - det(**A**) = $\prod\lambda_i$
- Odpowiedź:
  - wielomian stopnia n
    - w „sumie n+1 uporządkowanych składników" mamy:
      - pierwszy składnik: x^n
      - drugi składnik: $x^{n-1}(-x_1-x_2-...-x_n) = -x^{n-1}(x_1+x_2+...+x_n)$
      - ...
      - ostatni składnik: $x^0(-x_1)(-x_2)\ldots(-x_n) = (-1)^n x^0 x_1 x_2 \ldots x_n$

…

Rozważania nad macierzami 2x2

- Co spełniają pierwiastki wielomianu charakterystycznego?
  - z jednej strony
    - (zwykłe rozpisanie wzoru)
      $\det(\mathbf{A}-\mathbf{I}\lambda) = (a-\lambda)(d-\lambda) - bc = \lambda^2 - (a+d)\lambda^1 + (ad-bc)\lambda^0$
  - z drugiej strony:
    - zakładając, że znamy pierwiastki tego wielomianu, i że są to (niekoniecznie różne od siebie) λ_1 i λ_2
      $\det(\mathbf{A}-\mathbf{I}\lambda) = (\lambda-\lambda_1)(\lambda-\lambda_2) = \lambda^2 - (\lambda_1+\lambda_2)\lambda^1 + (\lambda_1\lambda_2)\lambda^0$
  - porównanie:
    - $\lambda^2 - (a+d)\lambda^1 + (ad-bc)\lambda^0$
    - $\lambda^2 - (\lambda_1+\lambda_2)\lambda^1 + (\lambda_1\lambda_2)\lambda^0$

Rozważania nad macierzami 2x2

- Niezależność kolumn macierzy a wyznacznik
  - wyznacznik macierzy 2x2

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

  - zależność (liniowa) kolumn wymaga, aby zachodziło

    $a/c = b/d$ (zał: $c \neq 0$ i $d \neq 0$)

    $ad = bc$

    $ad - bc = 0$

    $\det(\mathbf{A}) = 0$

Rozważania nad macierzami 2x2

- Niezależność kolumn macierzy a wyznacznik
  - (ogólnie) warunkiem zależności kolumn dowolnej macierzy **A** jest det(**A**) = 0
  - (w szczególności) warunkiem zależności kolumn macierzy **A**–**I**λ jest det(**A**–**I**λ) = 0

…

Dygresja

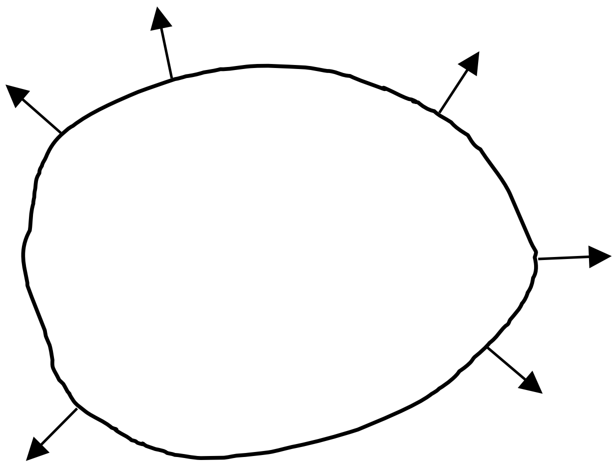
- Pomocne pojęcia
  - gradient
  - prostopadłość, ortogonalność
    - normalna krzywej
    - wektor normalny krzywej (strony prostej)
  - równoległość wektorów

- Wektory normalne w
  - ograniczeniach równościowych
  - funkcjach celu

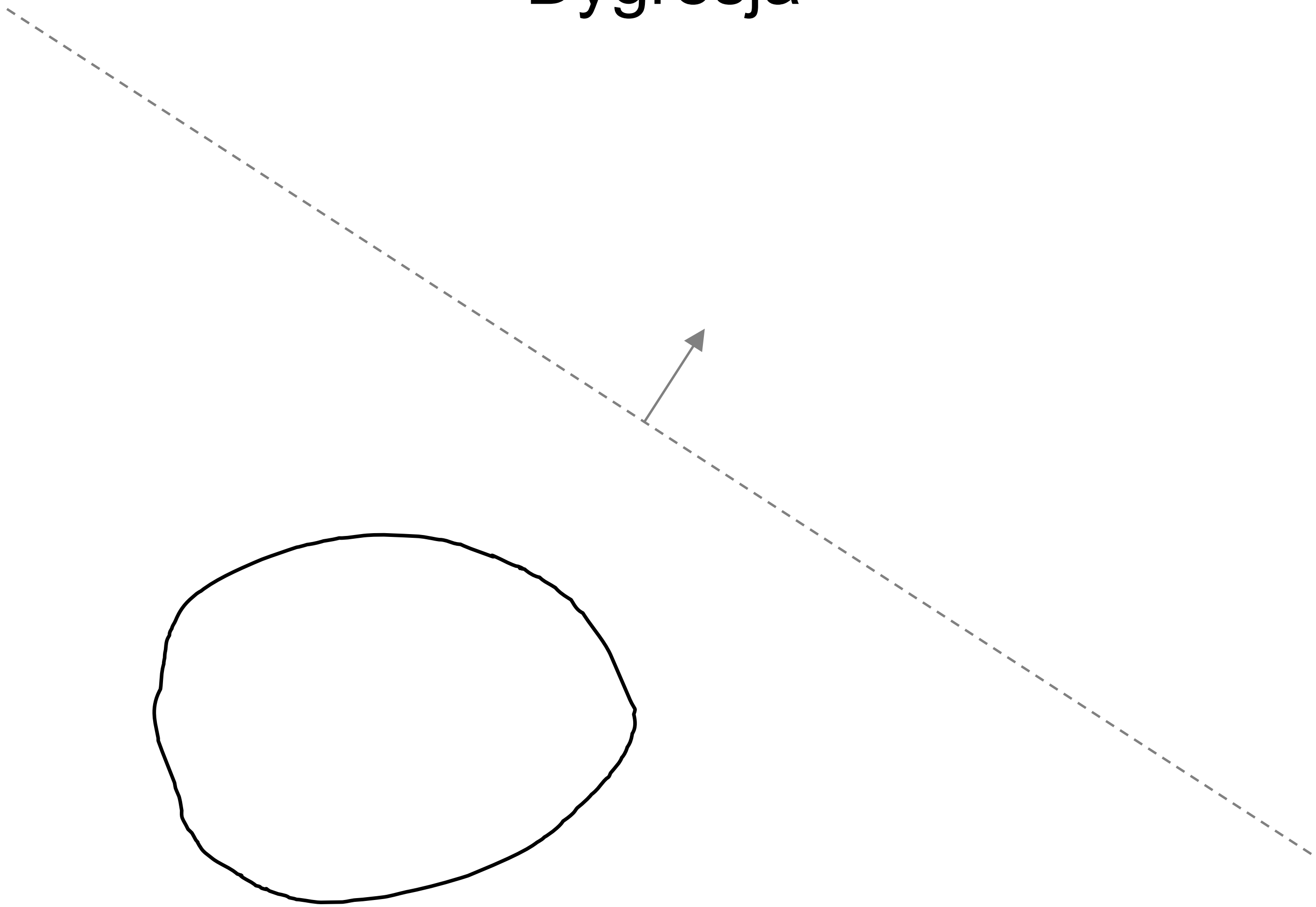
Dygresja

- Metoda mnożników Langrange’a

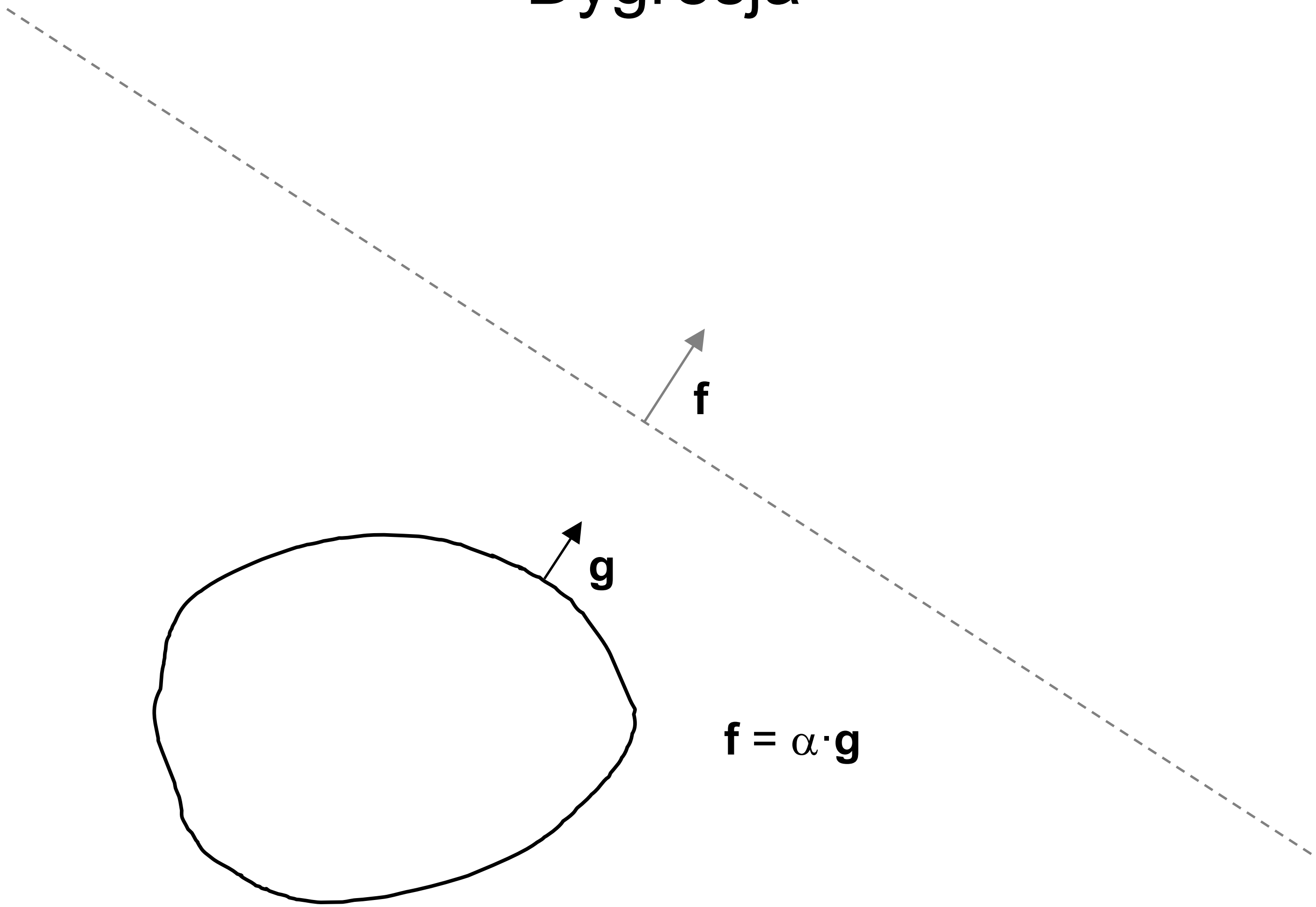
Dygresja



Dygresja



Dygresja



$\mathbf{f} = \alpha \cdot \mathbf{g}$

…

Dygresja

- Wpływ macierzy na „wytwarzane” przez nie wektory
  - długości
  - kierunki

Dygresja

- Wpływ macierzy na długości „wytwarzanych” przez nie wektorów
- Dokładniej
  - wpływ macierzy $\mathbf{A} = \mathbf{M}^T\mathbf{M}$ na długość wektora $\mathbf{Mx}$

    (powstałego z przemnożenia dowolnego wektora $\mathbf{x}$ przez /akurat/ macierz $\mathbf{M}$ taką, że $\mathbf{A} = \mathbf{M}^T\mathbf{M}$)

    - długość wektora $\mathbf{x}$: $(\mathbf{x}^T\mathbf{x})^{1/2}$
      - kwadrat długości wektora $\mathbf{x}$: $\mathbf{x}^T\mathbf{x}$
    - długość wektora $\mathbf{Mx}$: $((\mathbf{Mx})^T(\mathbf{Mx}))^{1/2} = (\mathbf{x}^T\mathbf{M}^T\mathbf{Mx})^{1/2} = (\mathbf{x}^T\mathbf{Ax})^{1/2}$
      - kwadrat długości wektora $\mathbf{Mx}$: $\mathbf{x}^T\mathbf{Ax}$

Dygresja

- Takie postawienie sprawy (tzn. zdefiniowanie powyższych wyrażeń) pozwala formułować konkretne problemy maksymalizacji/minimalizacji
  - np. interesować nas może maksymalna osiągana wartość $\mathbf{x}^T\mathbf{A}\mathbf{x}$
    - dla jakich (dowolnych) $\mathbf{x}$ wartość $\mathbf{x}^T\mathbf{A}\mathbf{x}$ jest maksymalna?
      - i ile wtedy wynosi?

Dygresja

- Zapisujemy

  max $\mathbf{x}^T\mathbf{A}\mathbf{x}$

  p.o. -brak-

  – inaczej

  max $\mathbf{x}^T\mathbf{A}\mathbf{x}$ / $\mathbf{x}^T\mathbf{x}$

  p.o. -brak-

  – inaczej

  max $\mathbf{x}^T\mathbf{A}\mathbf{x}$ / c^2

  p.o. $\mathbf{x}^T\mathbf{x} = c^2$ (c: constant)

  – inaczej

  max $\mathbf{x}^T\mathbf{A}\mathbf{x}$

  p.o. $\mathbf{x}^T\mathbf{x} = 1$ (powyższe dla c = 1)

Dygresja

- Ostateczna postać problemu maksymalizacji

  max $\mathbf{x}^T\mathbf{A}\mathbf{x}$

  p.o. $\mathbf{x}^T\mathbf{x} = 1$

- Rozwiązanie powyższego problemu (dla symetrycznej macierzy **A**) generuje maksymalną wartość własną tej macierzy i odpowiadający jej wektor własny
  - a po pewnych modyfikacjach: kolejne wartości/wektory własne

Dygresja

- Nowy kłopot:
  - jak rozwiązanie powyższego problemu ma się do faktu, że wartości własne macierzy są pierwiastkami pewnego wielomianu* a jej wektory własne są rozwiązaniami pewnych układów równań**?

    (* konkretnie: $\det(\mathbf{A} - \mathbf{I}\lambda) = 0$)

    (** konkretnie: $\mathbf{Ax} = \lambda\mathbf{x}$)

Dygresja

- (Nowe) rozwiązanie kłopotu
  - interpretacja rozwiązania rozważanego problemu „w duchu” metody mnożników Lagrange’a
  - niech
    - $F(\mathbf{x}) = \mathbf{x}^T\mathbf{A}\mathbf{x}$
    - $G(\mathbf{x}) = \mathbf{x}^T\mathbf{x} - 1$
  - wtedy problem maksymalizacji przyjmuje postać

    max $F(\mathbf{x})$

    p.o.: $G(\mathbf{x}) = 0$

Dygresja

- Dla tak postawionego problemu mamy
  - gradienty
    - $\partial F/\partial \mathbf{x} = 2\mathbf{Ax}$ (dla symetrycznej **A**)
    - $\partial G/\partial \mathbf{x} = 2\mathbf{x}$
  - przyrównanie gradientów (ze współczynnikiem λ)

$$\partial F/\partial \mathbf{x} = \lambda \cdot \partial G/\partial \mathbf{x}$$

$$2\mathbf{Ax} = \lambda \cdot 2\mathbf{x}$$

$$\mathbf{Ax} = \lambda \mathbf{x}$$

(λ jest wartością własną /gwarantującą nietrywialność rozwiązania/)

…