Matematika pro geometrickou morfometrii

Václav Krajíček

Vaclav.Krajicek@mff.cuni.cz

Department of Software and Computer Science Education
Faculty of Mathematics and Physics
Charles University

MFF UK
CGG
Prague

Přednáška 1

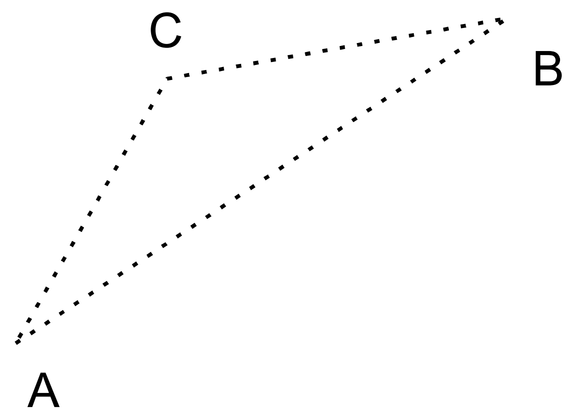
Zařazení do předmětu

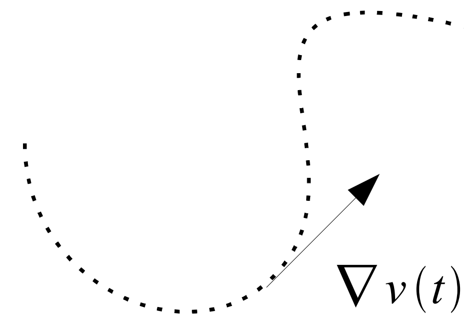
Anotace:

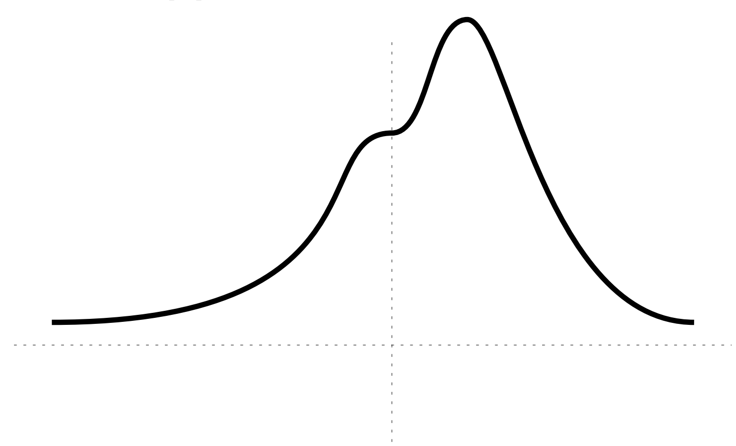
Předmět Biomedicínská metodologie II navazuje na předmět Biomedicínská metodologie I. Jeho cílem je vybavit studenty biomedicínské antropologie metodologickými znalostmi pro hodnocení morfologie objektů od přímého měření probandů až po digitální měření dvojrozměrných dat včetně jejich zpracování metodami klasické i geometrické morfometrie. Předmět je zaměřen na výuku metod, které analyzují hlavu od jeho vnějších oblastí až po vnitřní struktury, zaznamenávané nejčastěji pomocí RTG snímků, ale umožňující hodnotit také fotodokumentaci, výstupy PC tomografu či magnetické rezonance. Prvním cílem je zvládnout metodiku klasického měření hlavy, popř. jejich modelů v paměti počítače, včetně výpočtu proporčních indexů či základního zpracování dat metodami multivariační statistiky. Dalším cílem je představit a zároveň osvojit si vyhodnocování nejčastěji využívané zobrazovací techniky v medicíně. Studenti se naučí metricky vyhodnocovat dálkové RTG snímky hlavy pomocí speciálních softwarů, vyhodnocovat dentální věk dětí na základě analýzy panoramatických snímků a biologický věk dětí za použití RTG snímků ruky. Posledním cílem je představit metody geometrické morfometrie, včetně softwarů pro akvizici a statistické zpracování dvojrozměrných dat. Z tohoto hlediska na předmět dále úzce navazuje kurz 3D metod aplikovaných v antropologii.

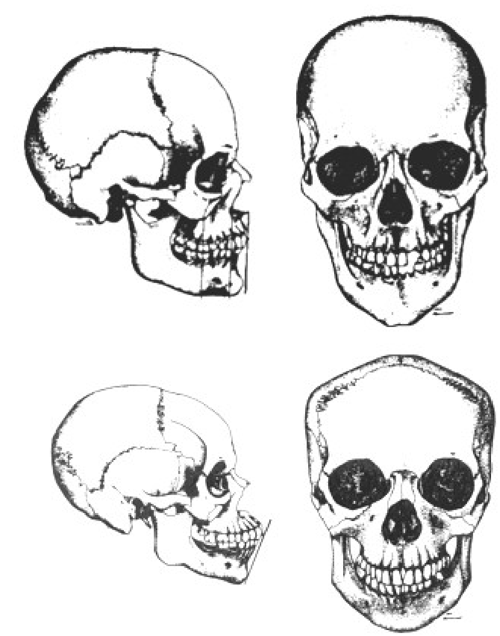
Úvod

- Rozhraní několika oborů
- Geometrie
- Morfometrie
- Biologie
- Antropologie
- Statistika









- Zpracování lékařských dat

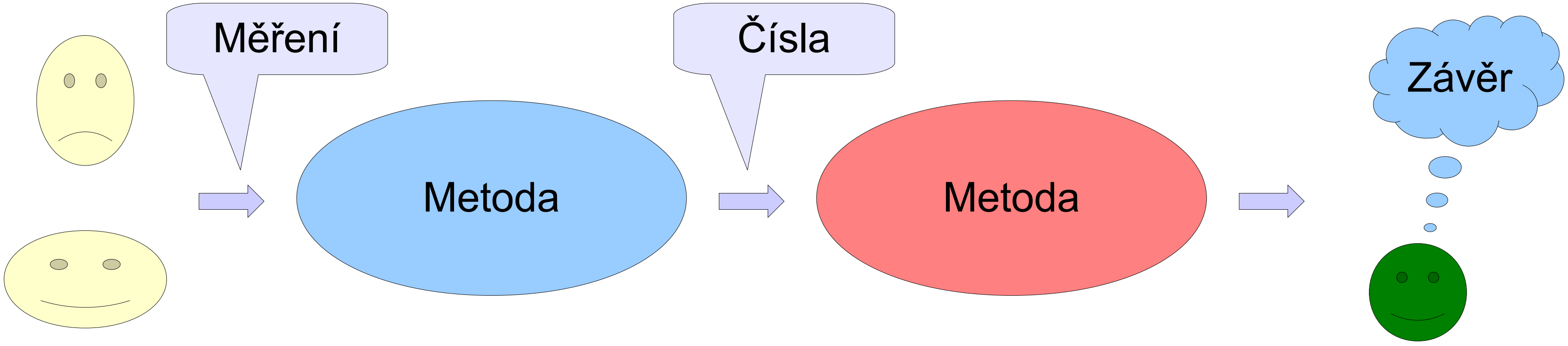
Cíle GMM

- Odhalit skrytá fakta a souvislosti
- Než začneme se vztahy → Popis exempláře
  - Landmarkové metody, křivky, trojúhelníkové sítě
- Nic nevíme → Získávání nových poznatků studiem tvaru u skupiny exemplářů.
  - PCA, Klastrová analýza, vizualizace
- Něco tušíme → Ověření nebo vyvrácení hypotézy.
  - Statistické testy, korelace, regresní analýza
- Víme co, nevíme kolik → Stanovení míry odlišnosti.
  - Metriky

Cíle GMM

(1) Popsat tvar čísly

(2) Čísla statisticky vyhodnotit



Lineární algebra
Geometrie

Statistika
Rozpoznávání vzorů

Cíle přednášky

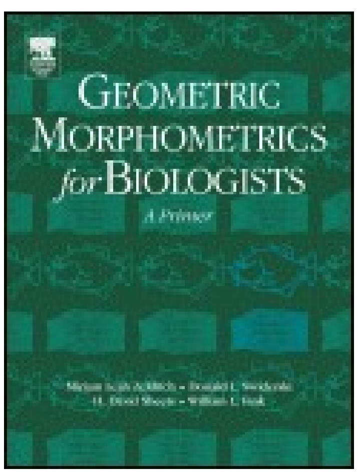
- Základy geometrické morfometrie s důrazem na pochopení matematických principů
  - Vzorečky a vztahy
- Názorné ukázky a příklady
  - Příklady s čísly
  - Matematický software, GMM software
- Naučit se používat vhodné metody pro danou úlohu
- Poskytnout příklady a zdroje inspirace
  - http://cgg.mff.cuni.cz/~vajicek/gmm

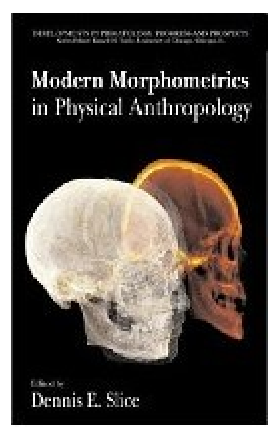
Osnova

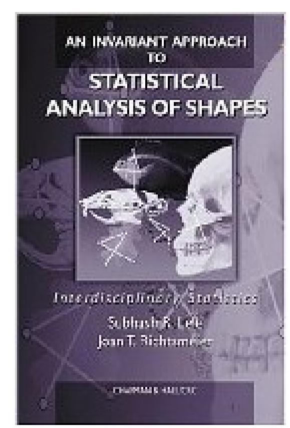
- Landmarkové metody
- Tvar
- Procrustovská analýza
- Shape space
- TPS, Warps
- PCA
- Nelandmarkové metody
- Křivky, transformace
- Statistika

Literatura

- Zelditch, Swiderski, Sheets, Fink: Geometric Morphometrics for Biologists: A Primer
- D.E. Slice: Modern Morphometrics in Physical Anthropology
- S. R. Lele, J. T. Richtsmeier: An Invariant Approach to Statistical Analysis of Shapes

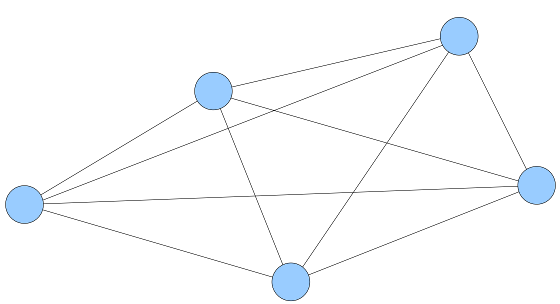




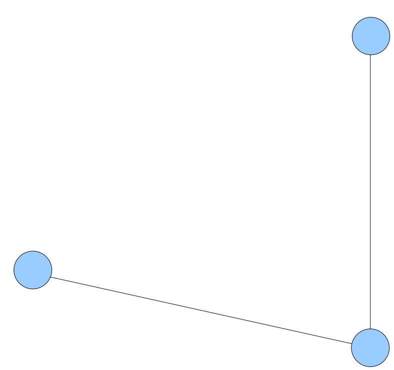


Tradiční metody

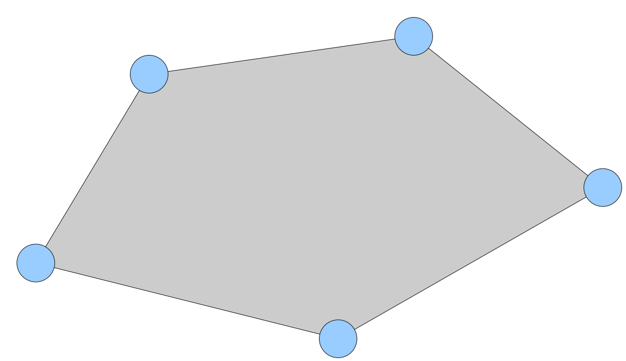
- Poloha „bodů“ nehraje roli (nedá se ani dobře měřit)
- Pracuje se s
  - Vzdálenost mezi dvěma „body“, poměry vzdáleností
  - Úhel mezi třemi „body“
  - (Plocha polygonu)



$$\binom{p}{2}$$



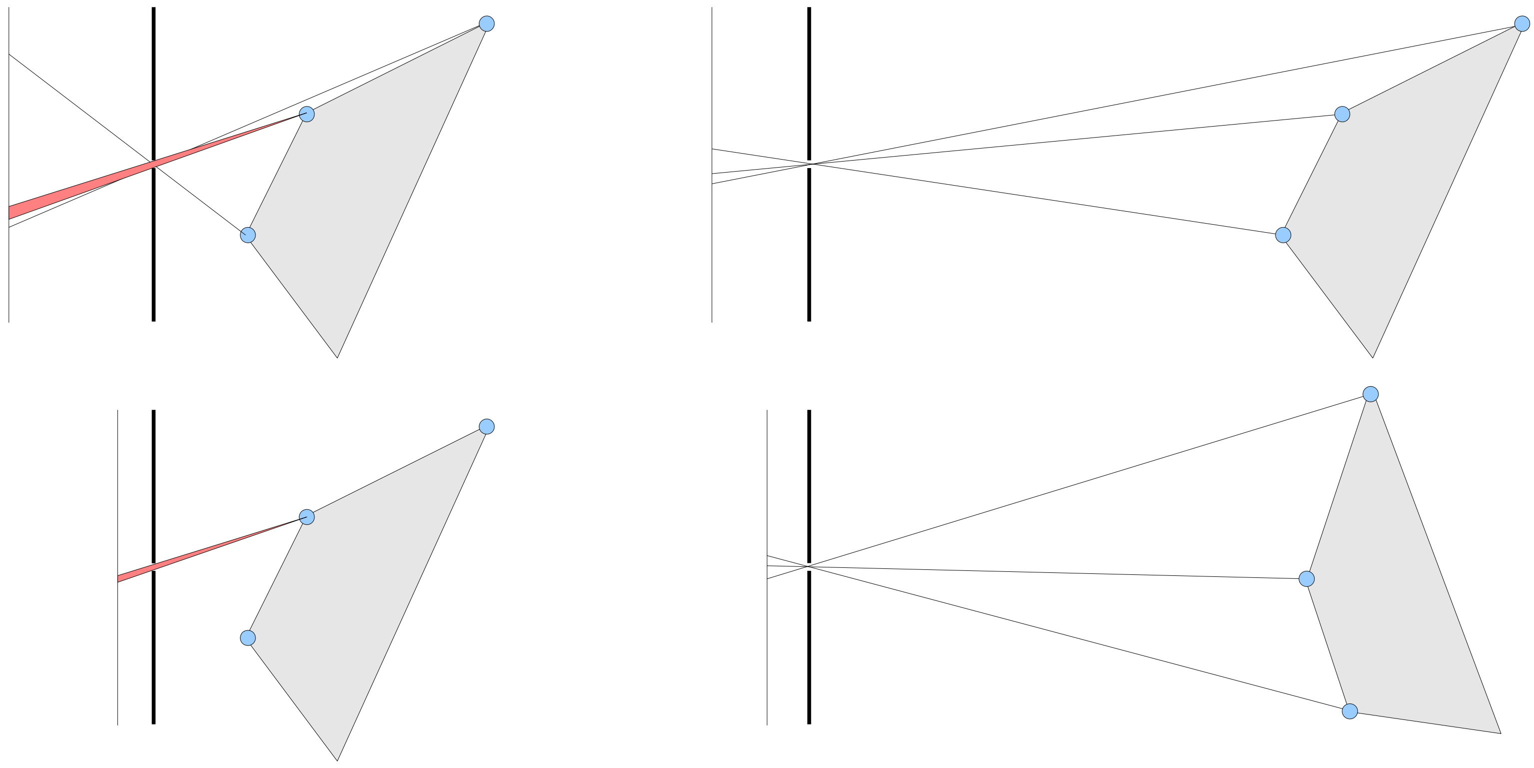
$$p\binom{p-1}{2}$$



Plocha složena z ploch nepřekrývajících se trojúhelníků

Získávání dat

- 2D → Fotografie, Kamera
  - Perspektivní projekce - potenciální příčina chyb měření
- 3D → Objemová data, povrchové skenery, polohovací zařízení



Landmark

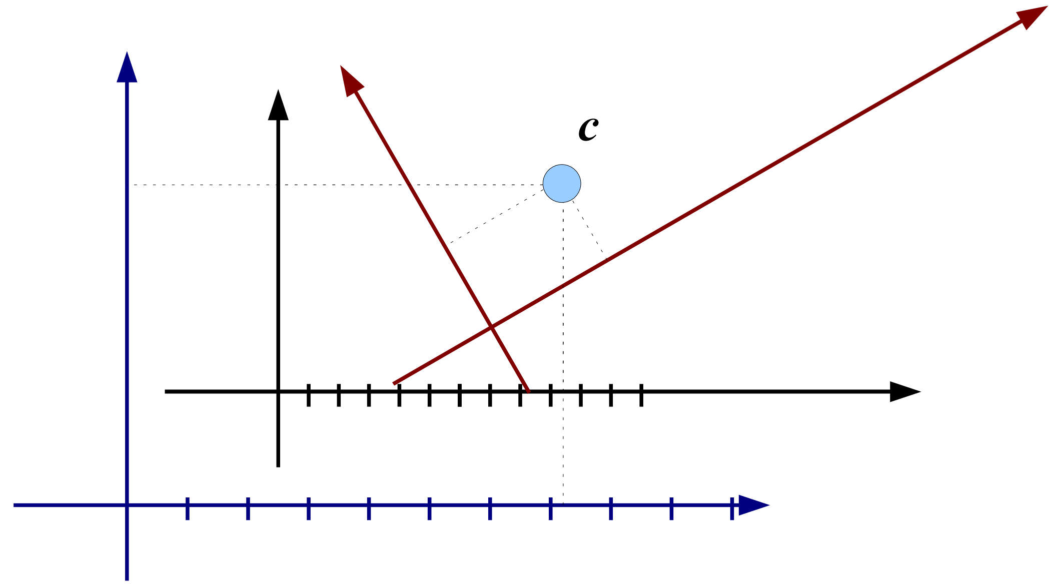
- Jméno – význam pro biologa
  - abstraktní, „nezávislé na reprezentaci“
  - 2D obrázek, exemplář, odlitek
- Souřadnice – význam pro matematika, algoritmus
  - Například dvojice nebo trojice „nějakých“ čísel
    - Dimenze, 3D, 2D
  - O jaky typ čísel jde?
    - Celá – rastrové obrázky, Desetinná – vektorový svět
    - Záporná/kladná – volba „počátku“
  - Co znamenají?
    - Vzdálenosti od „počátku“ v jednotlivých dimenzích

Výběr landmarků

- Závisí na tom co chceme najít
  - Pokud nevíme co hledáme, snažíme se pokrýt systematicky
- Pokud nás zajímá konkrétní část exempláře, musíme ji odpovídajícím způsobem pokrýt landmarky
- Pokud porovnáváme dvě části (např horní a dolní čelist) → pokrýt landmarky
- Někdy není pokrytí možné → jiné (nelandmarkové) metody
  - semi-landmarky
  - křivky

Jednoznačnost souřadnic

- Odpovídá každému landmarku právě jedna souřadnice?
- Volba počátku
- Volba jednotek
- Volba orientace
- Báze
- Lineární kombinace



$$\boldsymbol{c} = P + \sum k_i \boldsymbol{b}_i \qquad \boldsymbol{B} = \{\boldsymbol{b}_1, \boldsymbol{b}_2, \ldots\} \qquad \boldsymbol{k} = \{k_1, k_2, \ldots\}$$

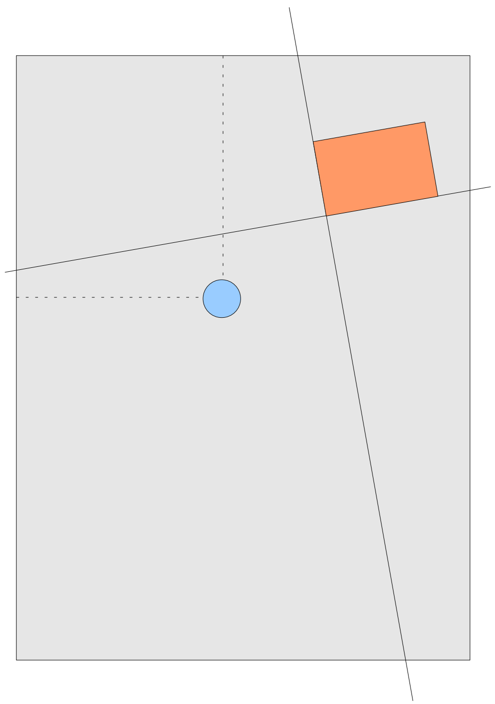
počátek

báze

souřadnice

Příklad

- Změřené kolmé vzdálenosti od okraje papíru
  - 3,53 cm od delší strany a 4,86 cm od kratší strany
  - v případě počítače může jít o pixely od okraje obrázku
- Víme, že objekt byl fotografován
  - z velké vzdálenosti
  - ne kolmo k rovině objektu
  - s kalibrační pomůckou ve scéně
- Kalibrace kamery
  - s pomocí lineární algebry



Vektory a skaláry

- Skaláry známe
- Vektory také známe → připomenutí
  - posloupnost čísel zachycující více-dimenzionální informaci
  - **v** = [1;2], **w** = [-3;2;-1], **u**=[1;3;4;5;6;12]
- Podobné operace jako se skaláry
  - [1;2] + [3;1] = [4;3]
- Vektorový, skalární součin

Matice

- Tabulka čísel

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \qquad \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \end{bmatrix} \qquad \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Obecnější případ vektorů

- Operace sčítání a násobení matic
- Musí souhlasit velikosti matic a vektorů

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} = \begin{bmatrix} \mathbf{30} & 36 & 42 \\ 66 & 81 & 96 \\ 102 & 126 & 150 \end{bmatrix}$$

- Symbolicky

$$A \; A = B$$

Transformace

- Práce se souřadnicemi lze snadno realizovat pomocí maticových operací
  - Posun, otočení, zvětšení množiny landmarků
- Proto se používá pro hromadné zpracování

$$\begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix} \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} = \begin{bmatrix} 30 \\ 36 \\ 42 \end{bmatrix}$$

Transformační matice — Souřadnice — Transformované souřadnice

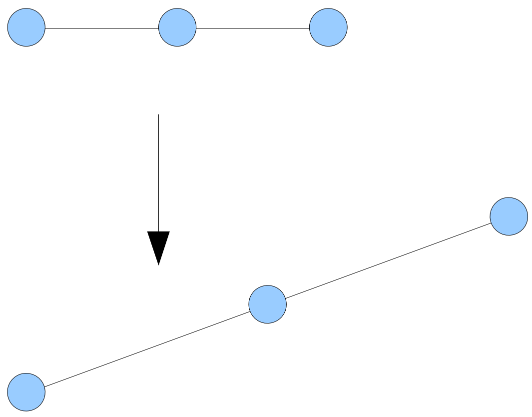
- Inverzní transformace

$$\begin{bmatrix} ? & ? & ? \\ ? & ? & ? \\ ? & ? & ? \end{bmatrix} \begin{bmatrix} 30 \\ 36 \\ 42 \end{bmatrix} = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}$$

Lineární transformace

- Linearita je důležitá vlastnost maticových transformací → nemění vzájemné vztahy mezi transformovanými vektory

$$\alpha A(\boldsymbol{a}+\boldsymbol{b})=\alpha A\boldsymbol{a}+\alpha A\boldsymbol{b}$$



- Užitečné transformace (otočení, změnšení/zvětšení) jsou lineární
- Posunutí se nedá realizovat násobením matic

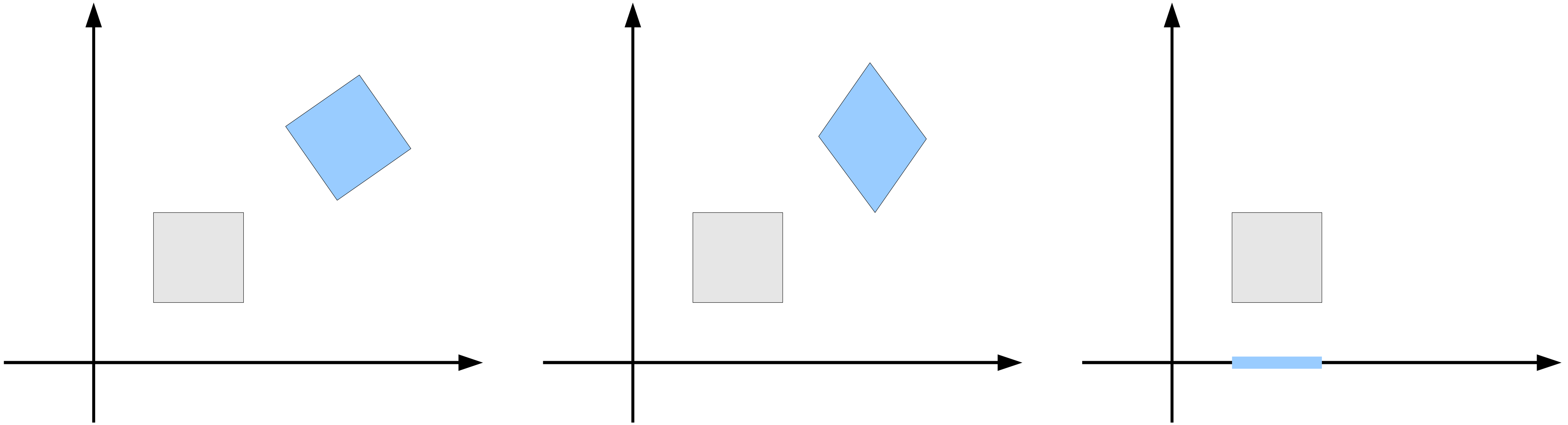
Homogenní souřadnice

- Perspektivní projekce není lineární transformace
  - nezachovává rovnoběžnost, úhly, vzdálenosti
- Lze ji popsat maticí za použití tzv. Homogenních souřadnic
  - Přechod do prostou vyšší dimenze (2D → 3D, 3D → 4D)
  - Z bodů se stávají přímky
  - V prostoru homogenních souřadnic transformace lineární
  - Převod do prostoru normálních souřadnic
- Naštěstí se s nimi často nesetkáme
  - software pro rekonstrukci 3D skenů
  - obecná kalibrace kamery

Typy transformací

- Podobnost
  - Otočení, posun, změna měřítka
- Afinní transformace
  - Podobnost + zkosení
- Projekce

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



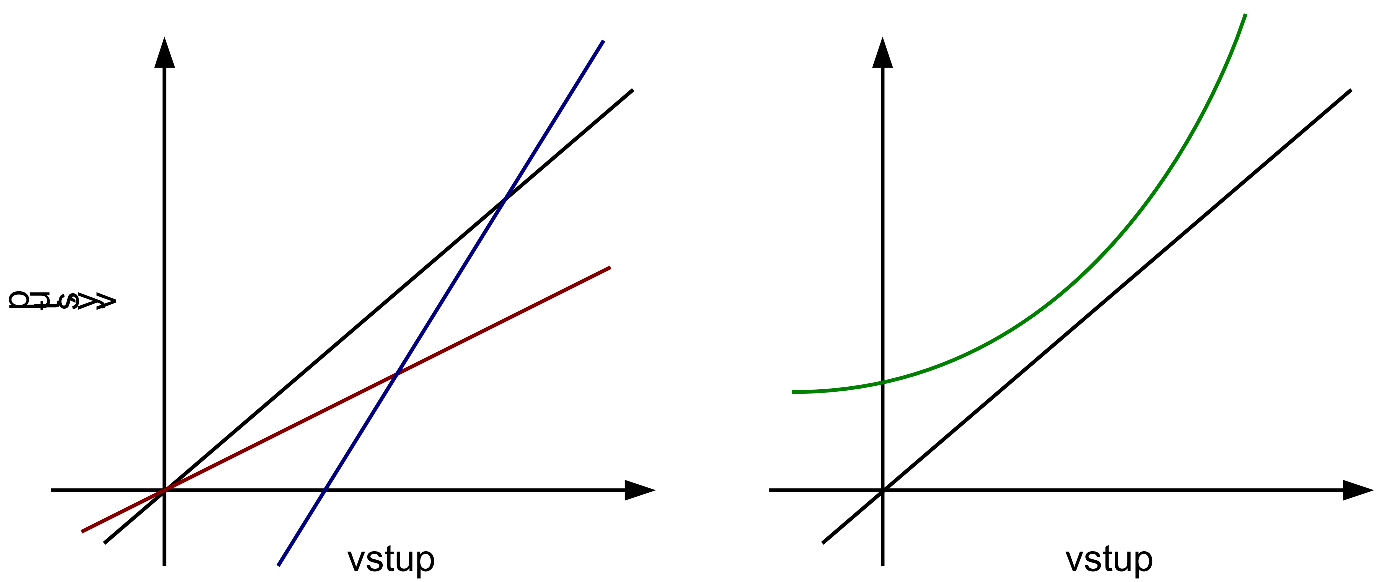
Maticová kalkulačka

- Praktická ukázka
- Software Matlab, Octave, Scilab

```
>> A=[1 0 0; 0 1 0; 0 0 1];
>> v=[1; 2; 3];
>> A*v
ans = 1 2 3
>> B=[cos(pi/4) -sin(pi/4); sin(pi/4) cos(pi/4)];
>> B*[0; 1]
ans = 0.7071 0.7071
```

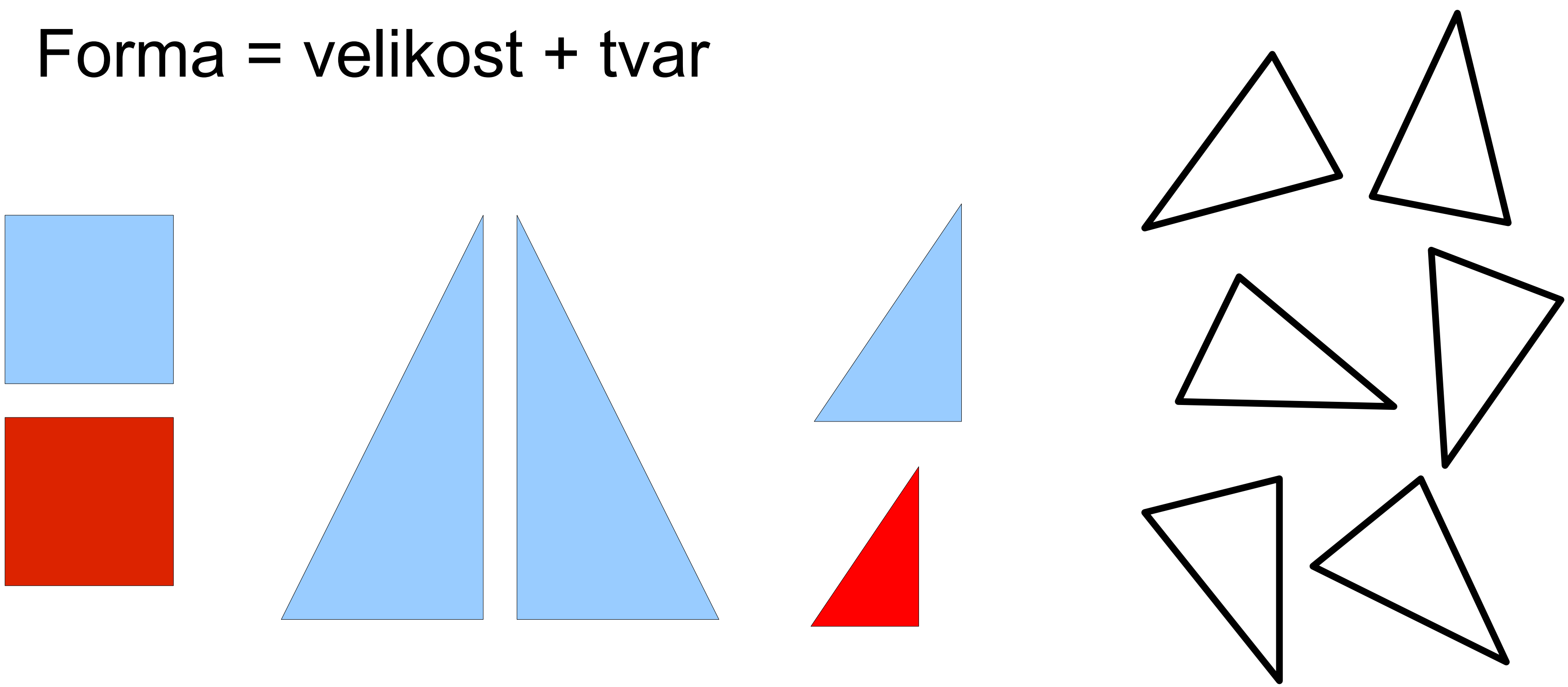
Nelineární transformace

- Zahrnují složitější operace než násobení koeficienty a sčítání
- Perspektivní projekce – zahrnuje dělení



Moderní metody GMM

- Metody měření, zpracování, analýzy a zobrazení s cílem studia tvaru a jeho změn
- Tvarové proměnné - popisují tvar, nemění se s velikostí
- Forma = velikost + tvar



Tvar

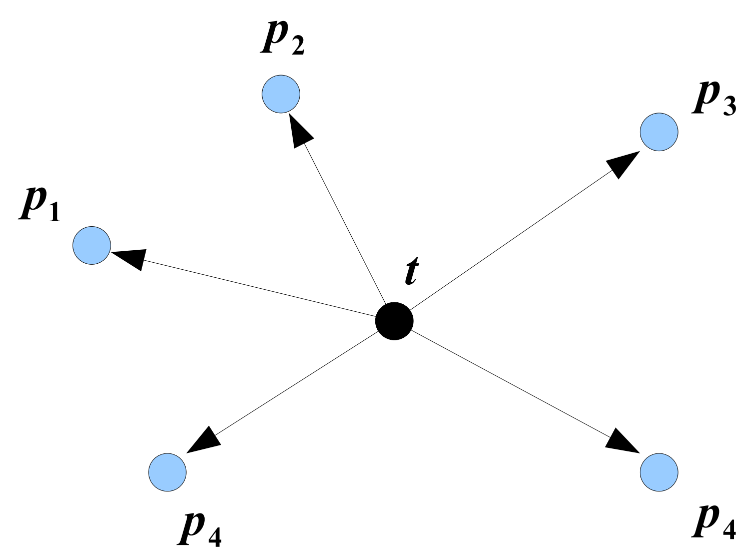
- Definice (Kendall)
  - Tvar je veškerá geometrická informace, která objektu zůstane po odstranění vlastností polohy, měřítka a otočení
- „Tradiční“ morfometrie
  - Pracuje přímo se vzdálenostmi
    - Vylepšení - poměry vzdáleností
  - Samotné vzdálenosti stěží odrážejí tvar
  - Dva exempláře jsou tak tvarem neporovnatelné
- Oddělení velikosti a tvaru v datech
  - Jedna z hlavních myšlenek geometrické morfometrie

Velikost

- Vlastnosti
  - Nezáporná, lineární
- Středová velikost (Centroid size)
  - Matematicky je „nezávislá na tvaru“
  - Výpočet z landmarků

$$\boldsymbol{t}=\frac{1}{n}\sum_{i=1}^{n}\boldsymbol{p}_i$$

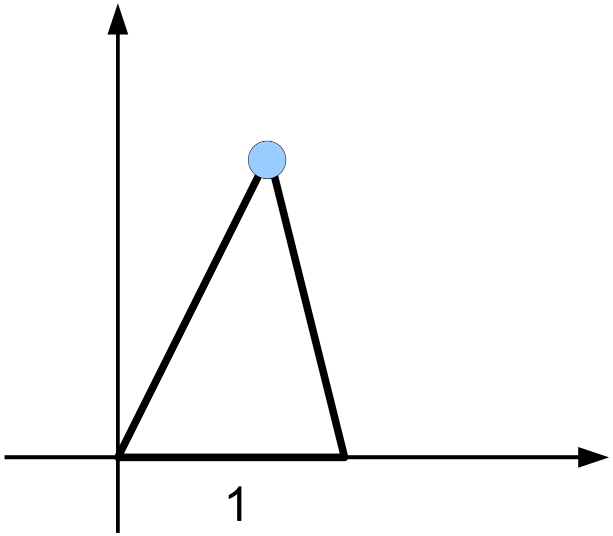
$$CS=\sqrt{\sum_{i=1}^{n}(\boldsymbol{p}_i-\boldsymbol{t})^T(\boldsymbol{p}_i-\boldsymbol{t})}$$



  - Často se hledá konkrétní souvislost středové velikosti s tvarem

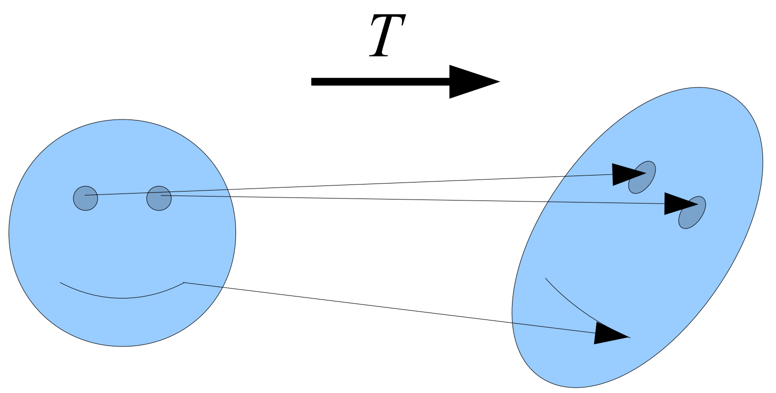
Příklad - trojúhelníky

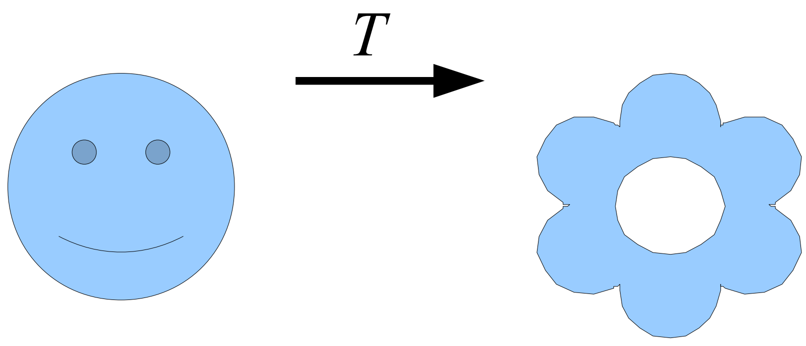
- Nejjednodušší geometrický objekt, u kterého se dá hovořit o tvaru
- K úplnému popisu tvaru trojúhelníku v rovině stačí dvojice čísel
- Stupeň volnosti tvaru = počet proměnných po odstranění velikosti, otočení a posunutí
  - Pro 2D $2\mathrm{p}-4$
  - Pro 3D $3\mathrm{p}-7$
  - Obecně $pk-k-k(k-1)/2-1$
    - k - dimenze, p - počet bodů



Registrace

- Definice pojmu
  - Hledání transformace dvou datových množin do společného systému souřadnic





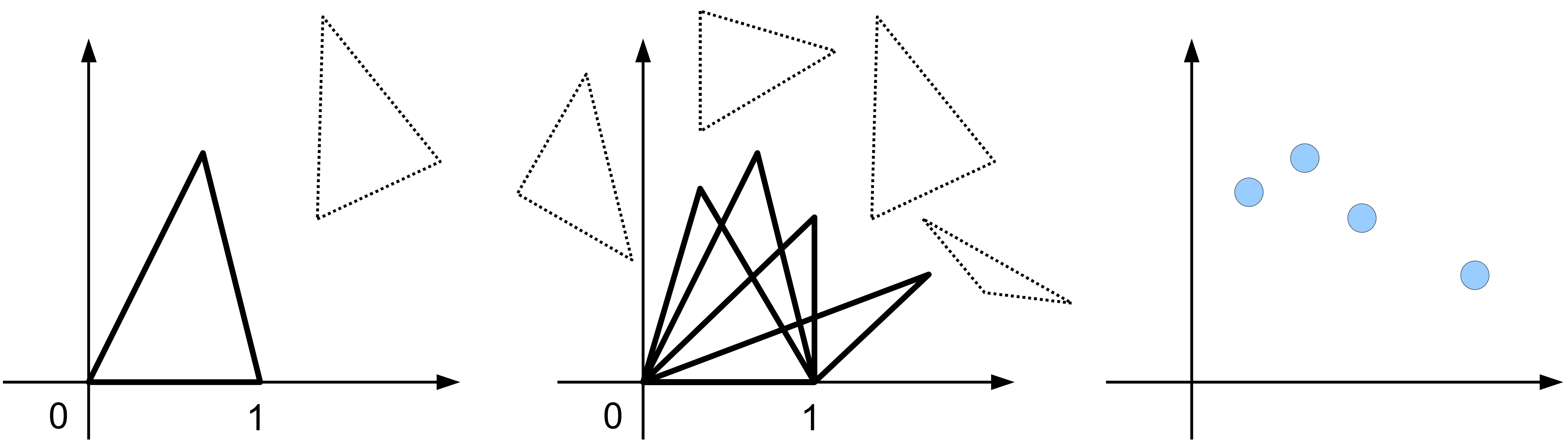
- Transformací vzorků do společného, „jednotkového“, souřadného systému odstraníme vliv velikosti na tvarové proměnné

Typy registrace

- Různé typy registrace podle
  - Dat se kterými pracují – 2D, 3D, objem, obrázek, body, trojúhelníkové sítě
  - Typu transformace které připouštějí – rigidní, elastická
  - Způsobu hledání – uzavřené formy, gradientní metody
  - Způsobu porovnání dvou datových množin
- Pro účely získání tvarových proměnných připustíme pouze transformaci polohy, orientace a měřítka
  - Dá se popsat lineární algebrou (matice)

Dvoubodová registrace

- Transformované souřadnice → Booksteinovy tvarové proměnné, „Booksteinova registrace“
- Volba dvojice landmarků, které budou tvořit tzv. základnu (base-line)
- Transformace celého vzorku tak, aby základna ležela na ose x mezi 0 a 1.



Dvoubodová registrace

- Naměřené souřadnice vzorku → matice X
  - p1 a p2 tvoří osu

$$X = \begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \\ x_3 & y_3 \end{bmatrix}$$

- Posun jednoho z vrcholů báze do počátku

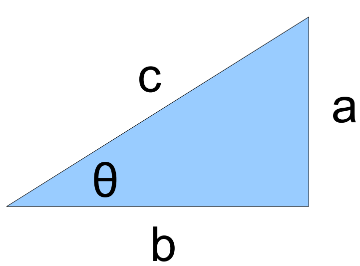
$$X' = X - \mathbf{1}^T \boldsymbol{p}_1$$

- Výpočet úhlu který svírá báze a osa x

$$\cos\theta = b/c$$

$$c = \sqrt{x_2'^2 + y_2'^2}, b = x_2'$$

$$\theta = \cos^{-1} x_2' / \sqrt{x_2'^2 + y_2'^2}$$



Dvoubodová registrace

- Rotace do osy x

$$X'' = HX' \qquad H = \begin{bmatrix} \cos(-\theta) & \sin(-\theta) \\ -\sin(-\theta) & \cos(-\theta) \end{bmatrix}$$

- Škálování na délku 1

$$X''' = X''/x_2''$$

- Aplikuje se na všechny vzorky
- Sníží se počet proměnných o 4 → získáme stejný počet jako je statistických tvarových proměnných
- Závislé na výběru základny
- V 3D se vybírá základní trojúhelník

Praktická ukázka

- Bookstein transformace v Octave/Matlab
- Past
- Tvarové parametry jsou souřadnice třetího vrcholu

$$\theta = \cos^{-1} x_2' / \sqrt{x_2'^2 + y_2'^2}$$

```
>> D=[3 1; 1 5; 8 3];

>> CD=D-ones(3,1)*D(1,:);

>> theta=acos(CD(2,1) / sqrt( CD(2,1)^2 + CD(2,2)^2) );

>> H=[cos(-theta) -sin(-theta); sin(-theta) cos(-theta) ];

>> RCD=(H*CD')'

>> SRCD=RCD./RCD(2,1)

SRCD =   0      0      1.0000     0     -0.1000     -1.2000
```

Klouzající základna

- Podobně jako předchozí algoritmus
- Škálování celého objektu na jednotkovou velikost
- Do počátku se umístí střed základny