Analýza hlavných komponentov

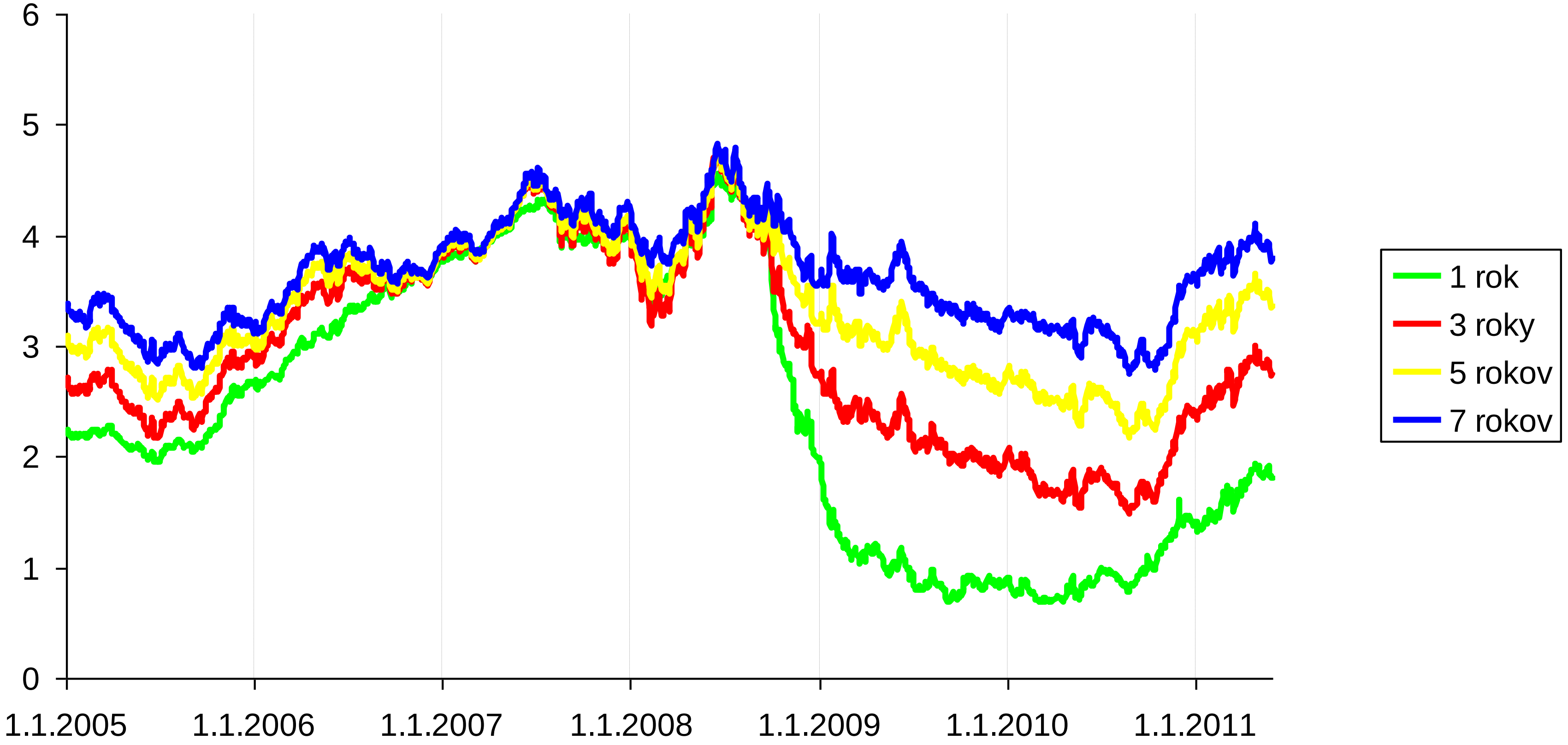
Motivácia

- Úloha: Navrhnite scenáre zmien výnosovej krivky pre účely stresového testovania v dlhopisovom portfóliu

- Problém: Výnosová krivka sa skladá z väčšieho počtu bodov, v ktorých je vývoj čiastočne korelovaný

Motivácia

- Čo hovoria údaje:

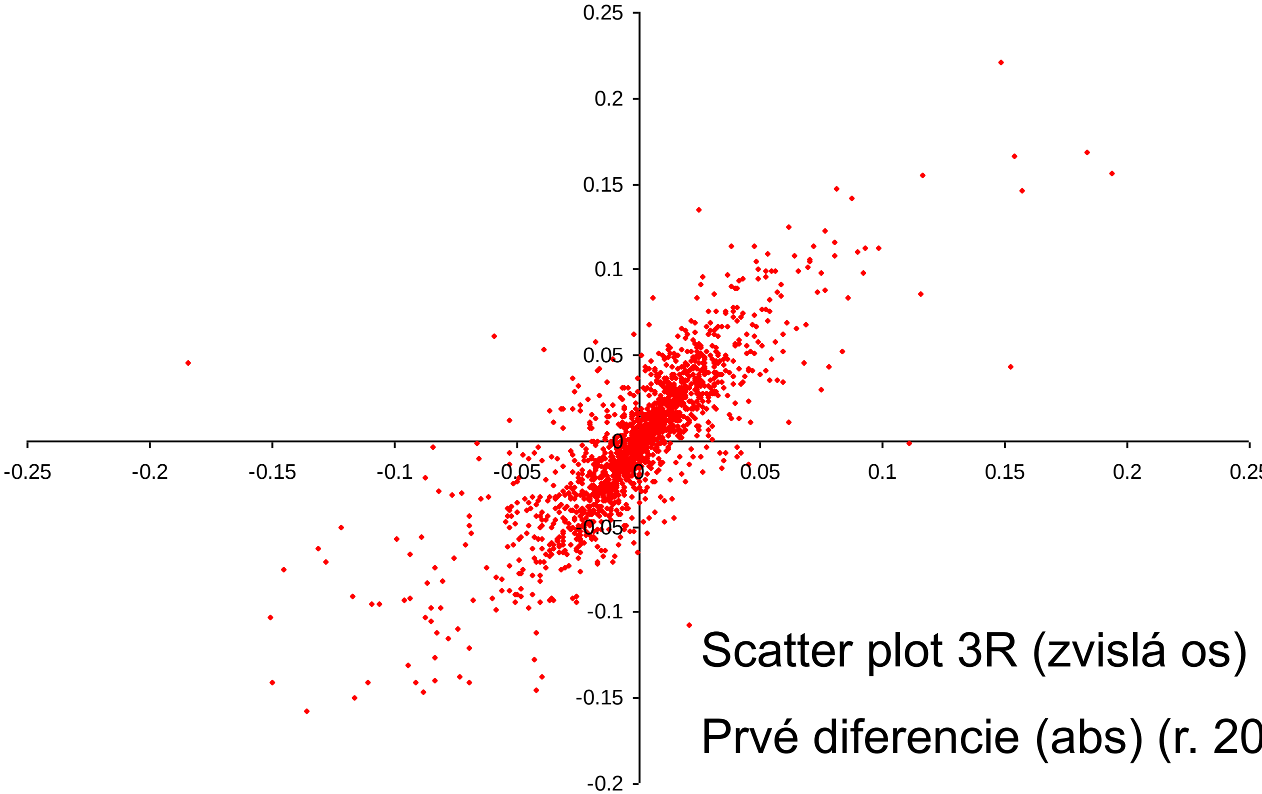


Motivácia

- Čo hovorí regulácia:
  - Zákon o bankách (§ 33f):
    - Ekonomická hodnota banky nesmie klesnúť o viac než 20 % z hodnoty vlastných zdrojov banky následkom náhlej a neočakávanej zmeny úrokových mier na trhu.
  - Opatrenie NBS o rizikách a systéme riadenia rizík (§ 5):
    - Náhlou a neočakávanou zmenou úrokových mier na trhu sa rozumie paralelný posun výnosovej krivky smerom nahor o 200 bázických bodov.

Základná myšlienka

- Ak by boli úrokové sadzby iba dve:
  - elipsu možno reprezentovať jej hlavnou osou

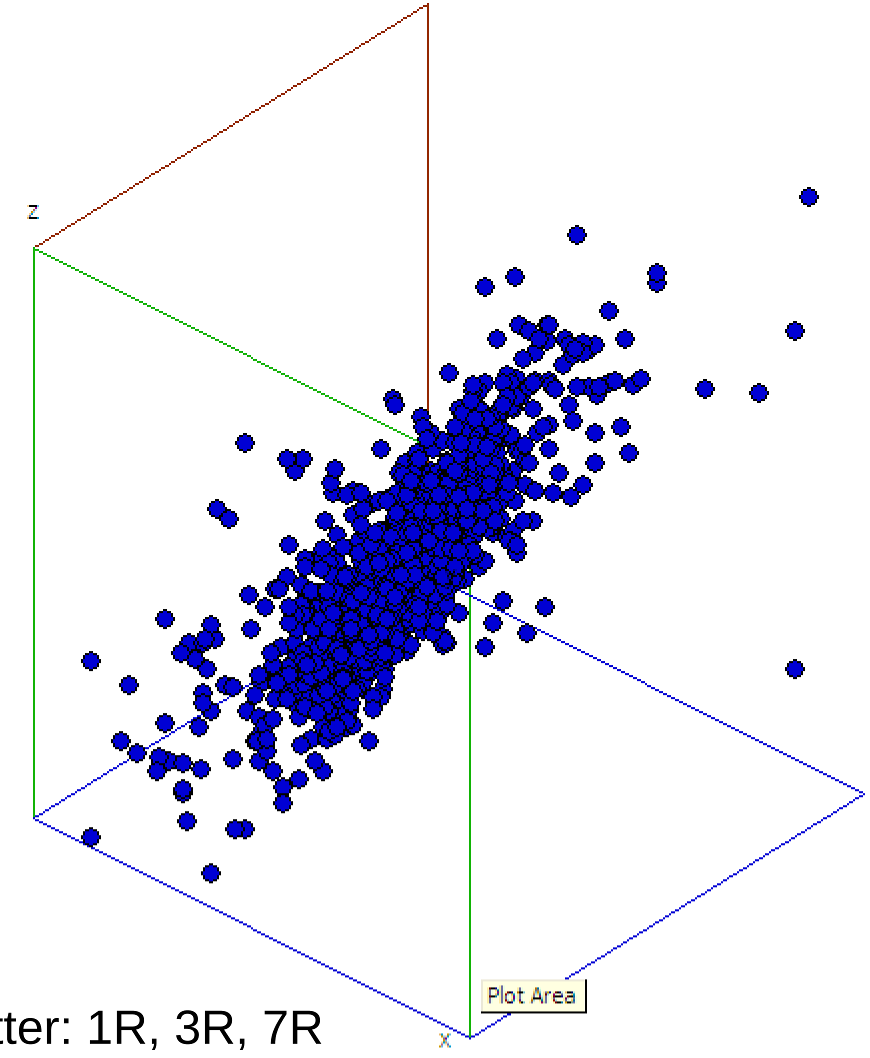


Scatter plot 3R (zvislá os) os 1R,

Prvé diferencie (abs) (r. 2005 – 6/2011)

Základná myšlienka

- Ak by boli úrokové sadzby tri:
  - Elipsoid možno stlačiť do elipsy



3D scatter: 1R, 3R, 7R

PCA: hrubý náčrt

- PCA = analýza hlavných komponentov (principal component analysis)

- V manažmente rizík máme často problém s veľkým počtom rizikových faktorov
  - akciové portfólio s veľkým počtom akcií
  - výnosová krivka daná veľkým počtom bodov
  - pozície v rôznych menách

- Cieľom je identifikovať najvýznamnejšie typické pohyby rizikových faktorov
  - systémový pohľad na riziko
  - využitie pri návrhu stresových scenárov

- Základná myšlienka: Znížiť dimenziu!

- Cieľom je nájsť malý počet lineárnych kombinácií pôvodných časových radov
  - mali by vysvetľovať čo najväčšiu časť variability pôvodných údajov
  - mali by byť navzájom nekorelované

Teoretické východiská

- Predpoklad: X je $N \times T$ matica zmien rizikových faktorov (časové rady sú v riadkoch)

- Východisko: Spektrálna dekompozícia matice:

  Každú symetrickú maticu A možno zapísať v tvare

  $$A = P \Lambda P^T,$$

  kde Λ je diagonálna matica vlastných hodnôt matice A a
  P je ortonormálna matica štandardizovaných vlastných vektorov matice A (stĺpce matice P).

- Aplikácia dekompozície na kovariančnú maticu Σ pôvodných údajov: $\Sigma = P \Lambda P^T$

- Výpočet $N \times T$ matice hlavných komponentov:

  $$Y = P^T (X - \mu)$$

Vlastnosti HK

- Platí
  - $E(Y) = 0$
  - $\mathrm{cov}(Y) = P^T \Sigma P = P^T P \Lambda P^T P = \Lambda$

- Dôsledok: HK sú nekorelované a ich variancie sú rovné vlastným číslam

- Transformácia na HK predstavuje centrovanie a rotáciu

- Keďže celková variabilita pôvodných premenných je

$$\sum_{i=1}^{n} \mathrm{var}(X_i) = \mathrm{trace}(\Sigma) = \sum_{i=1}^{n} \lambda_i$$

tak časť celkovej variancie vysvetlená j-tym HK je $\dfrac{\lambda_j}{\sum_{i=1}^{n} \lambda_i}$

Vlastnosti HK

- Prvý HK je štandardizovaná lineárna kombinácia *X*, ktorá má maximálnu varianciu spomedzi všetkých takýchto kombinácií
  - Smer najväčšej variancie *N*-rozmerného oblaku pôvodných údajov

- Výberové hlavné komponenty:
  - Ak *X* má normálne rozdelenie a cov(*X*) má všetky vlastné čísla rôzne, tak výberové HK sú odhadmi skutočných HK náhodného vektora *X*, ak sa na odhad cov(*X*) použije ML.

- Získané HK nie sú invariantné na zmenu škálovania premenných
  - PCA sa aplikuje na štandardizované premenné, resp. na korelačnú maticu

- Výpočet kovariancie *X* a *Y*: $\mathrm{Cov}(X_i, Y_j) = p_{ij}\, \lambda_j$. Odvodenie?
- Ako vypočítať koreláciu?

Určenie počtu HK

- Vlastné hodnoty vyššie ako priemer (Kaiser)
- Stanoviť limit na percento vysvetlenej variancie
- Zlom v grafe vlastných čísiel (scree plot)

Interpretácia PCA

- Komponentové váhy (component loadings):
  - Ide o štandardizované vlastné vektory
  - Koeficienty predstavujú váhu pôvodných premenných pri tvorbe HK
  - Možno ich získať aj pomocou OLS

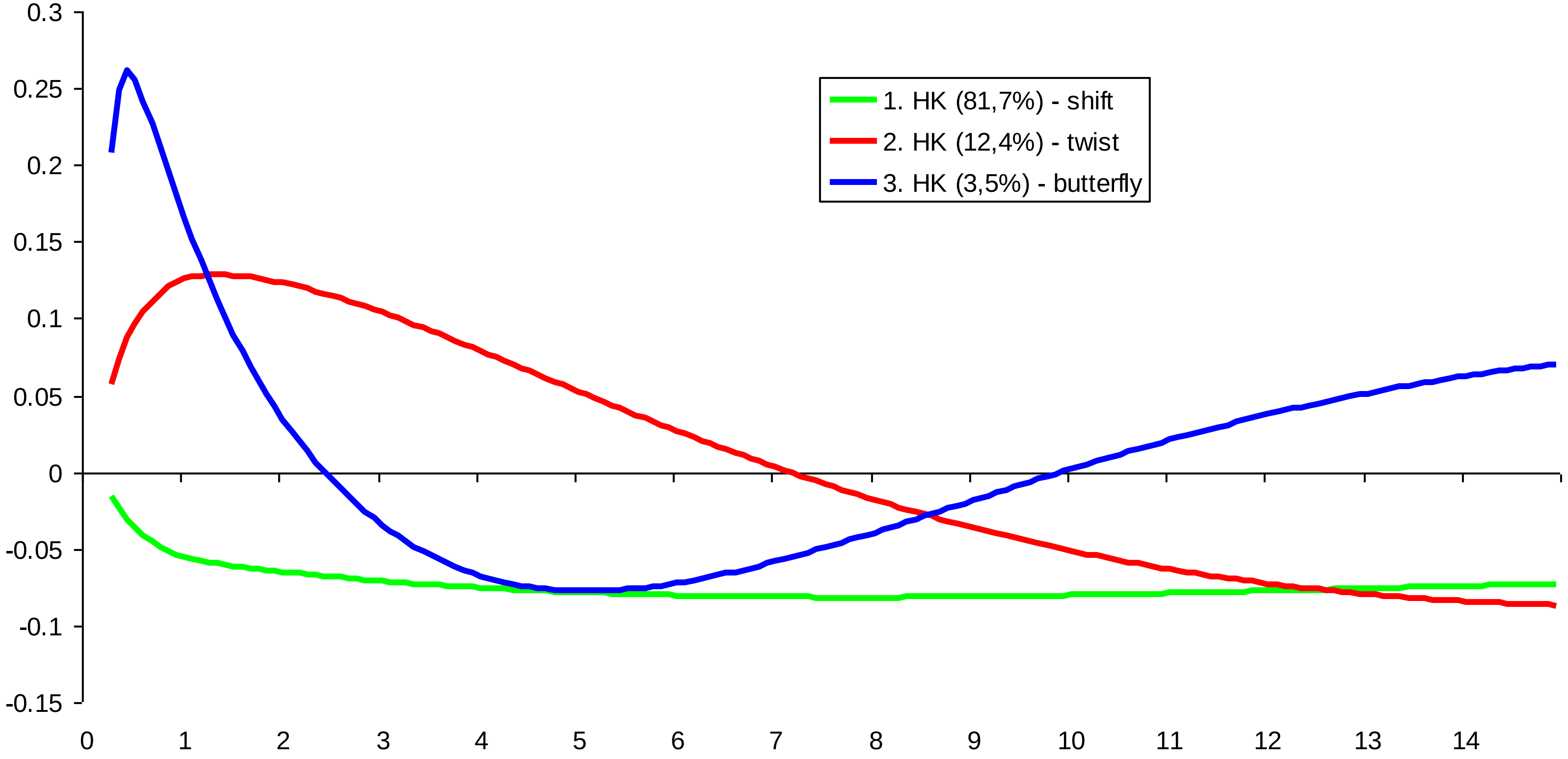
- Komponentové skóre (component scores)
  - Súradnice objektov v priestore HK

- Biplot = zobrazenie objektov v priestore HK

- Čo by sa stalo, ak by sme pôvodné premenné necentrovali?

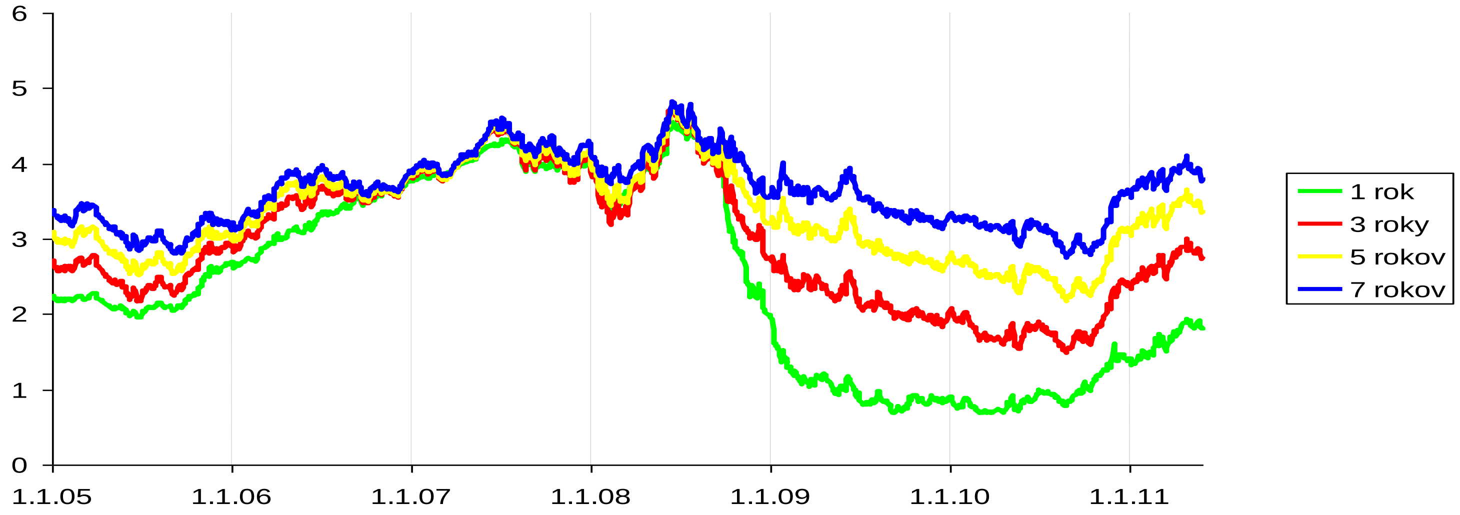
Výsledky

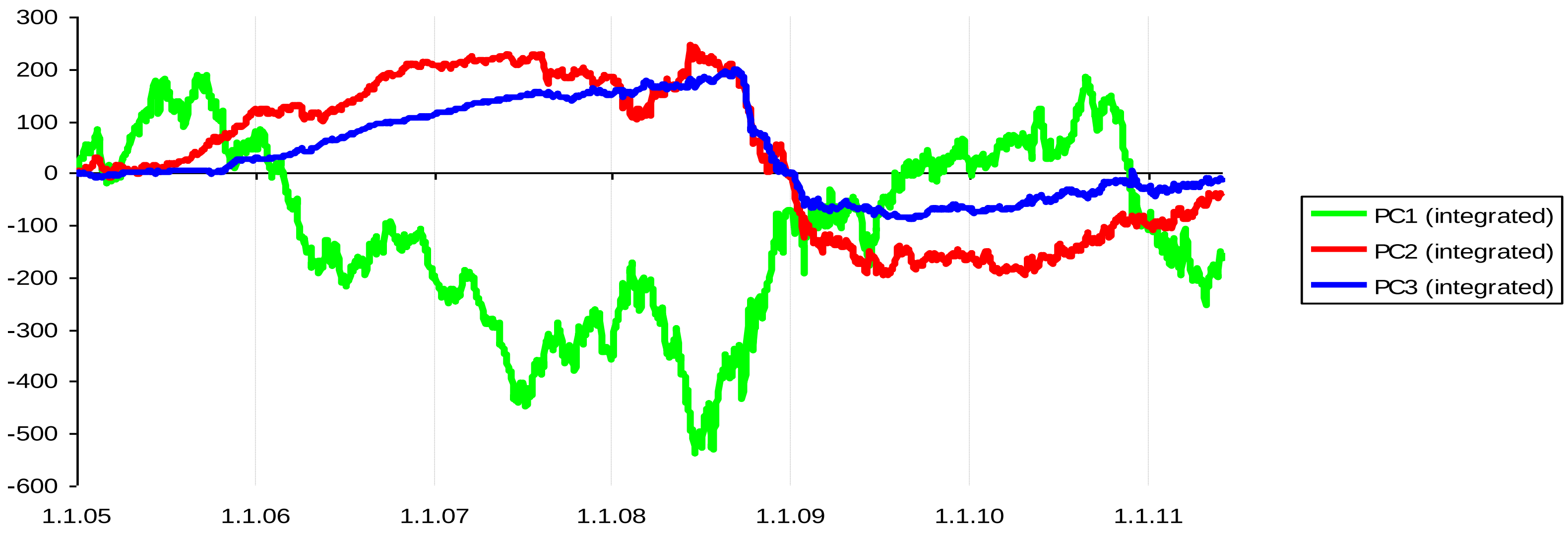
- Časové rady: denné absolútne zmeny (prvé diferencie) úrokových sadzieb
- Component loadings – prvé 3 HK (97,7%)



Výsledky

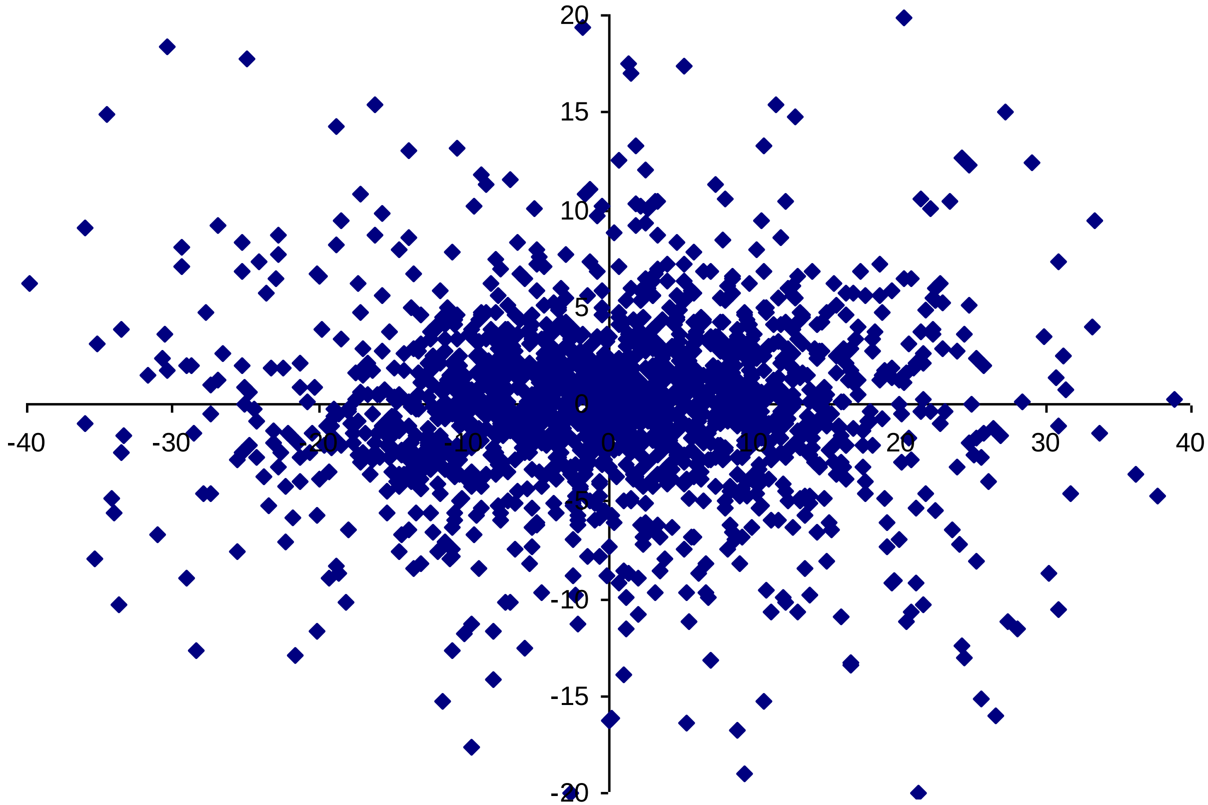
- Component scores





Výsledky

- Ako identifikovať, ktorý typ posunu prevažoval v jednotlivé dni?



- Scatter plot – zobrazenie v priestore (PC1, PC2)

Veľkosť scenárov

- Poznáme kvalitatívny opis „najreprezentatívnejších“ pohybov úrokovej krivky

- Ako určíme veľkosť príslušných scenárov?

Veľkosť scenárov

Príklad pre 1. HK

- Podmienené normálne rozdelenie k danému dňu: N(μ_{HK}, $\sigma^2{}_{HK}$)

- Zvolíme $y = (\mu_{HK} - u_\alpha \sigma_{HK}, 0, ..., 0)^T$

- Prenesenie zmeny spätnou transformáciou:

$$\mathrm{x} = Py + \mu$$

- Využívame, že
  - $P^T = P^{-1}$, kde stĺpce P tvoria vlastné vektory matice Σ
  - HK sú navzájom lineárne nezávislé

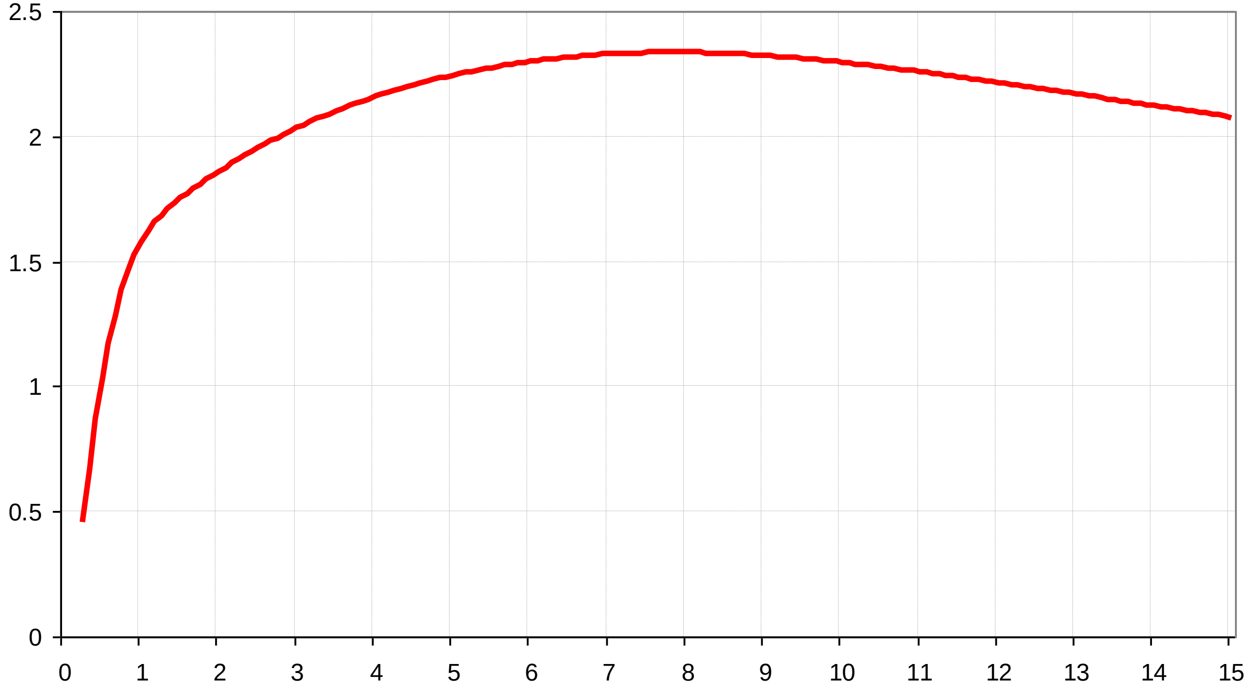
- Po zjednodušení:

$$\mathrm{x} = P_1\ (\mu_{HK} - u_\alpha \sigma_{HK}) + \mu,$$

kde P_1 je vlastný vektor prislúchajúci najvyššej vlastnej hodnote.

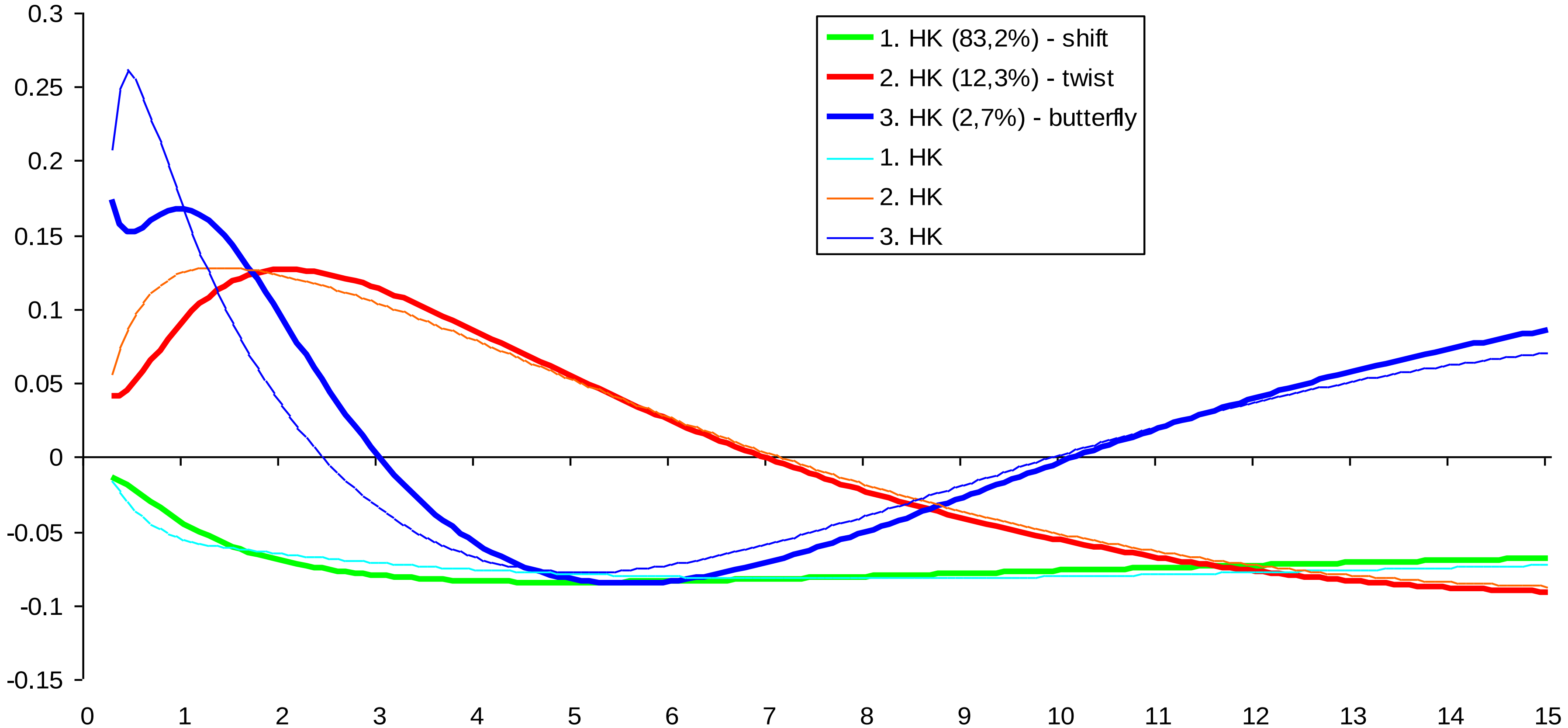
Veľkosť scenárov

- Výsledok pre „najtypickejší“ scenár s pravdepodobnosťou 99%:



Prípad kovariančnej matice

- Component loadings – prvé 3 HK (98,3%)



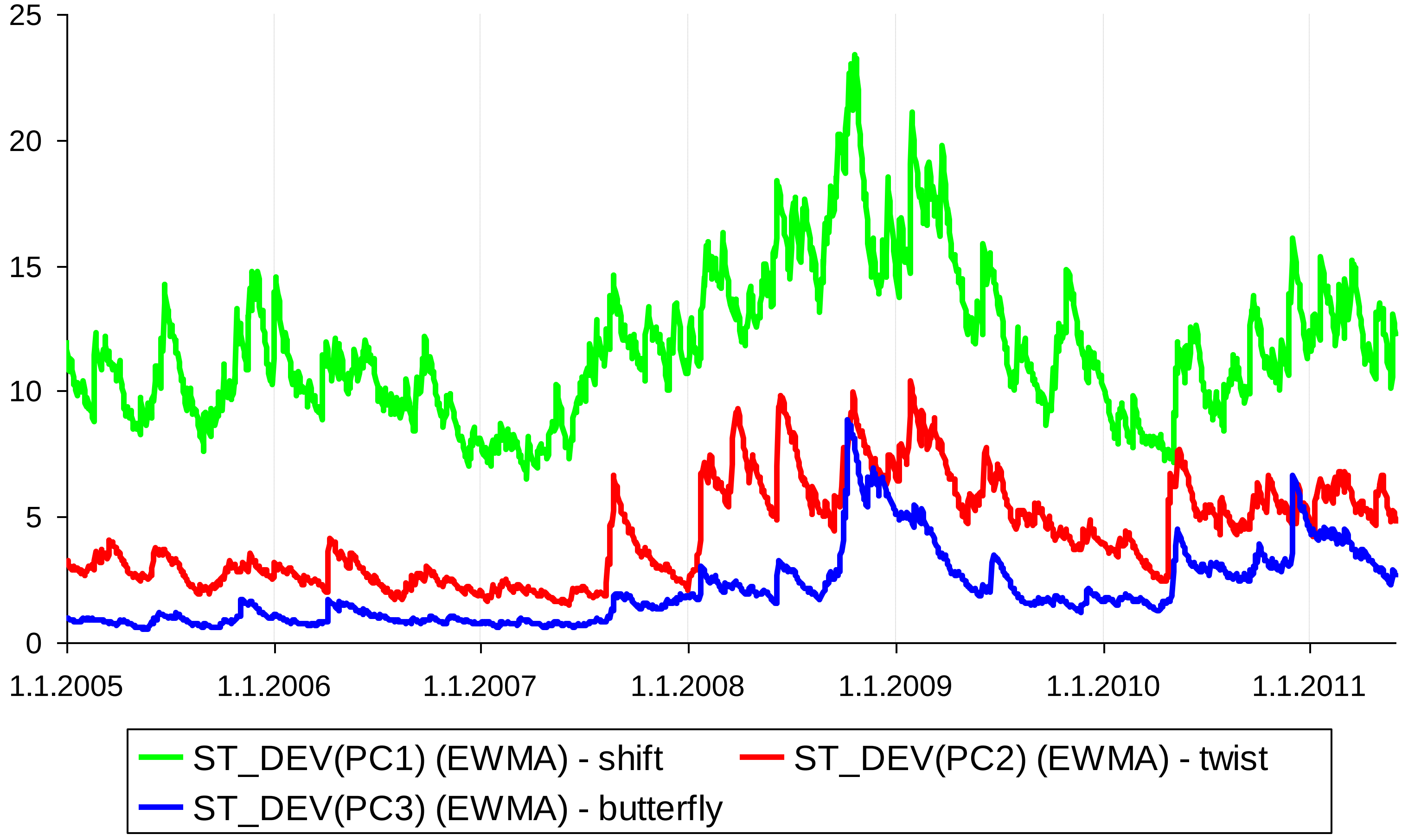
- Slabými čiarami sú zobrazené predchádzajúce výsledky získané z korelačnej matice

Kritika

- Hlavné komponenty často nie sú dobre interpretovateľné
- Pokrytá je iba lineárna závislosť (t.j. korelácia)
  - Nezachytáva fakt, že závislosť môže byť v čase trhových turbulencií vyššia
- Neumožňuje časovo premenlivé váhy pri konštrukcii HK

Kritika

- Ilustrácia, že faktory sú nekorelované, ale istý typ závislosti existuje



Možné rozšírenia

1. Využitie podmienenej kovariančnej / korelačnej matice
2. Orthogonal GARCH
3. PCA ako medzivýpočet v riadení rizík
4. Faktorová analýza – rotácia faktorov s cieľom zvýšiť interpretáciu