Pruned Skewed Kalman filter and Smoother: with Application to the Yield Curve and Asymmetric Monetary Policy Shocks

Gaygysyz Guljanov, Willi Mutschler, Mark Trede

30.08.2023

Introduction

Introduction: Contributions

1. Pruned skewed Kalman filter

2. Smoothing step with skewed Kalman filter

3. Real data application:
    - Dynamic Nelson-Siegel
    - Monetary DSGE model

- Replication codes are in:
    1. `https://github.com/wmutschl/pruned-skewed-kalman-paper`
    2. `https://github.com/gguljanov/pruned-skewed-kalman`

- Smets-Wouters model [Smets and Wouters, 2007] estimation in the follow-up paper

Introduction: New filter

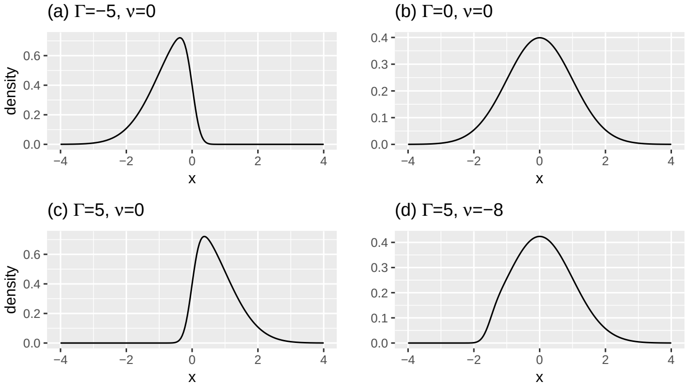
- ☐ Normal $\Rightarrow$ CSN
- ☐ Gaussian Kalman filter $\Rightarrow$ Pruned skewed Kalman filter
- ☐ Gaussianity as a special case
- ☐ CSN:
  - location parameters (μ, Σ)
  - skewness parameters (ν, Γ, Δ)

CSN distribution

CSN distribution

PDF:

$$\underbrace{\frac{1}{\Phi_q(0;\nu,\Delta+\Gamma\Sigma\Gamma')}}_{\text{CDF}}\underbrace{\phi_p(x;\mu,\Sigma)}_{\text{PDF}}\underbrace{\Phi_q(\Gamma(x-\mu);\nu,\Delta)}_{\text{CDF}},\quad x\in\mathbb{R}^p$$



CSN distribution

$$X = \Big(W | Z \geq 0\Big) \sim \text{CSN}(\mu,\ \Sigma,\ \Gamma,\ \nu,\ \Delta)$$

CSN properties:

- Linear transformation: $(A \cdot X)$ $\rightsquigarrow$ a problem
- Conditioning: $(X_1 | X_2)$
- CSN + Normal $\sim$ CSN: $(X + Y)$
- CSN + CSN $\sim$ CSN: $(X + Z)$ $\rightsquigarrow$ a problem

$\Rightarrow$ enable closed-form recursions for skewed Kalman filter.

- ☐ Joint distribution: $(X_1; X_2)^T$

Summation property

$$X \sim CSN_{p,q_x}(\mu_x, \Sigma_x, \Gamma_x, \nu_x, \Delta_x)$$

$$Y \sim CSN_{p,q_y}(\mu_y, \Sigma_y, \Gamma_y, \nu_y, \Delta_y)$$

$$Z = X + Y \sim CSN_{p,q_z}(\mu_z, \Sigma_z, \Gamma_z, \nu_z, \Delta_z)$$

$$q_z = q_x + q_y$$

Skewed Kalman filter

Skewed Kalman filter

Model:

$$X_t = BX_{t-1} + \eta_t, \qquad \eta_t \sim CSN(\mu_\eta, \Sigma_\eta, \Gamma_\eta, \nu_\eta, \Delta_\eta) \qquad \rightsquigarrow \text{State}$$

$$Y_t = AX_t + \epsilon_t, \qquad \epsilon_t \sim N(0, \Sigma_\epsilon) \qquad \rightsquigarrow \text{Observation}$$

Prediction step:

$$\left(X_t | Y_{1:t-1}\right) = B \times \left(X_{t-1} | Y_{1:t-1}\right) + \eta_t$$

$$X_t | Y_{1:t-1} \sim CSN(\mu_{t|t-1}, \Sigma_{t|t-1}, \Gamma_{t|t-1}, \nu_{t|t-1}, \Delta_{t|t-1})$$

Problem:

- $\Gamma_{t|t-1}, \nu_{t|t-1}, \Delta_{t|t-1}$ keep increasing in dimension

Reducing dimensions

Reducing dimensions

$$(W|Z \geq 0) = X \sim \mathrm{CSN}_{p,q}(\mu, \Sigma, \Gamma, \nu, \Delta)$$

$$\begin{pmatrix} W \\ Z \end{pmatrix} \sim N_{p+q}\left(\begin{bmatrix} \mu \\ -\nu \end{bmatrix}, \begin{bmatrix} \Sigma & \Sigma\Gamma' \\ \Gamma\Sigma & \Delta + \Gamma\Sigma\Gamma' \end{bmatrix}\right)$$

No correlation between W and Z

$\Rightarrow$ $(W|Z \geq 0)$ is normal

$\Rightarrow$ Prune weakly correlated elements

Reducing dimensions: Example

- Random variable X:

$$X \sim CSN\left(0, 1, \begin{bmatrix} 6 \\ 0.1 \end{bmatrix}, \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} 1 & -0.1 \\ -0.1 & 1 \end{bmatrix}\right)$$

- Correlation matrix of $\begin{pmatrix} W \\ Z \end{pmatrix}$:

$$R = \begin{pmatrix} 1.0000 & 0.9864 & 0.0995 \\ 0.9864 & 1.0000 & 0.0981 \\ 0.0995 & 0.0981 & 1.0000 \end{pmatrix}$$

- Random variable X after pruning dimensions:

$$\tilde{X} \sim \text{CSN}(0, 1, 6, 0, 1)$$

Reducing dimensions: Example

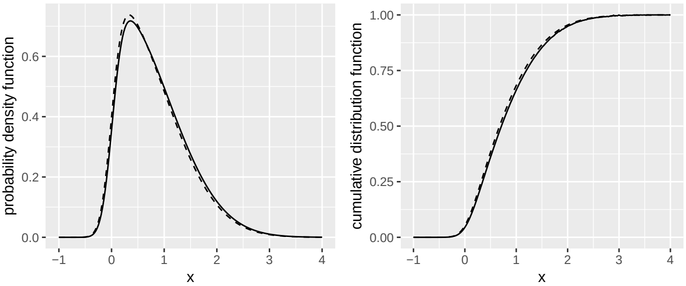


Figure 1: Solid lines: CSN with two skewness dimensions
Dashed lines: CSN with one skewness dimension

Reducing dimensions: Fade away

- Are correlations always sufficiently small?
- Dimensions fade away over time
- Intuition:

$$\begin{pmatrix} \Sigma & \Sigma\Gamma' \\ \Gamma\Sigma & \Delta + \Gamma\Sigma\Gamma' \end{pmatrix} \quad \longleftrightarrow \quad \frac{Cov_{ij}}{\sqrt{Var_i}\sqrt{Var_j}}$$

  - Σ converges to a constant matrix
  - Elements of $\Gamma\Sigma$ decrease over time
  - Diagonal elements of $\Delta + \Gamma\Sigma\Gamma'$ increase over time
  - Correlations decrease over time
- In other words, the algorithm is guaranteed to reduce the skewness dimension after sufficiently many periods

Smoothing

- We are first to derive smoothing step for skewed Kalman filter
  - approach as in Chiplunkar and Huang [2021]
- Pruning approach makes smoother computationally feasible

Monte Carlo study

Monte Carlo study

Accuracy:

1. The difference, between PSKF and GKF, is
   - rather small in the univariate case
   - really measurable in multivariate case
2. PSKF is numerically almost equivalent to non-pruned
3. Pruning threshold
   - does not make measurable difference in univariate case
   - only a small numericall difference in multivariate case

Monte Carlo study

Speed:

- PSKF is only mildly slower (8ms vs 1ms) than GKF
- Threshold of 10^{-2} is at least twice faster than 10^{-5}

Monte Carlo study

ML estimation of skewness parameters:

- Good
- GKF is biased, about mean and variance
- PSKF recovers skewness, and also gaussianity

Yield curve estimation

Yield curve estimation, Diebold and Li [2006]

$$y_t(\tau) = L_t + S_t \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} \right) + C_t \left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau} \right)$$

$$\begin{pmatrix} L_t - \mu^L \\ S_t - \mu^S \\ C_t - \mu^C \end{pmatrix} = \begin{pmatrix} G_{11} & G_{12} & G_{13} \\ G_{21} & G_{22} & G_{23} \\ G_{31} & G_{32} & G_{33} \end{pmatrix} \begin{pmatrix} L_{t-1} - \mu^L \\ S_{t-1} - \mu^S \\ C_{t-1} - \mu^C \end{pmatrix} + \begin{pmatrix} \eta_t^L \\ \eta_t^S \\ \eta_t^C \end{pmatrix}$$

- *Level* (L_t), *Scope* (S_t) and *Curvature* (C_t)
- λ: exponential decay rate
- τ: maturity
- The same dataset as Diebold and Li [2006]
  - Data: January 1985 - December 2000
  - 17 maturities
  - 39 parameter to be estimated

Yield curve estimation

	KF	$PSKF$	KF	$PSKF$	KF	$PSKF$	KF	$PSKF$
	L_{t-1}		S_{t-1}		C_{t-1}		μ	
L_t	0.9957 (0.008)	1.0004 (0.009)	0.0285 (0.009)	0.0253 (0.009)	−0.0222 (0.011)	−0.0218 (0.011)	8.2506 (1.086)	6.5516 (3.445)
S_t	−0.0303 (0.016)	−0.0015 (0.014)	0.9385 (0.018)	0.9767 (0.019)	0.0395 (0.021)	0.0399 (0.020)	−1.3786 (0.499)	−1.3411 (0.925)
C_t	0.0244 (0.023)	0.0085 (0.024)	0.0232 (0.026)	−0.0005 (0.027)	0.8428 (0.031)	0.8491 (0.030)	−0.3647 (0.383)	−0.3324 (0.476)

Table 1: Parameter estimates of G, μ^L, μ^S, and μ^C. Left side of each double column corresponds to estimates obtained with the conventional Kalman filter (*KF*), right side to estimates obtained with the pruned skewed Kalman filter (*PSKF*). Asymptotic standard errors appear in parentheses.

Yield curve estimation results

	L_t KF	L_t PSKF	S_t KF	S_t PSKF	C_t KF	C_t PSKF	Γ_η KF	Γ_η PSKF
L_t	0.0948 (0.008)	0.1906 (0.046)	−0.0140 (0.011)	−0.0668 (0.052)	0.0436 (0.019)	0.1648 (0.105)	0	−3.4648 (0.683)
S_t			0.3823 (0.030)	0.7546 (0.115)	0.0092 (0.034)	0.0565 (0.142)	0	−1.9895 (0.244)
C_t					0.8019 (0.081)	1.6045 (0.354)	0	1.2147 (0.225)

Table 2: Parameter estimates of Σ_η and Γ_η. Left side of a double column corresponds to estimates obtained with the conventional Kalman filter (KF), right side to estimates obtained with the pruned skewed Kalman filter (*PSKF*). Asymptotic standard errors appear in parenthesis.

Yield curve estimation

$$\widehat{COV}[\eta_t]^{KF} = \begin{pmatrix} 0.0948 & -0.0140 & 0.0436 \\ -0.0140 & 0.3823 & 0.0092 \\ 0.0436 & 0.0092 & 0.8019 \end{pmatrix}$$

$$\widehat{COV}[\eta_t]^{PSKF} = \begin{pmatrix} 0.0943 & -0.0181 & 0.0453 \\ -0.0181 & 0.3716 & 0.0223 \\ 0.0453 & 0.0223 & 0.8076 \end{pmatrix}$$

- Likelihood ratio test:
  test-statistic = 28.86, p-value = 2.396366e-06
- Hints to misspecification of the gaussian model

Yield curve estimation

	Decay λ	Standard deviation of measurement error for maturity 3	6	9	12	15	18	21	24
KF	0.07776 (0.002)	26.83 (8.68)	7.55 (3.66)	9.03 (2.85)	10.45 (3.11)	9.91 (2.96)	8.65 (2.65)	7.86 (2.45)	7.21 (2.24)
PSKF	0.07783 (0.002)	26.54 (8.51)	7.35 (3.57)	9.11 (2.85)	10.48 (3.11)	9.93 (2.96)	8.65 (2.65)	7.85 (2.45)	7.19 (2.23)

	Standard deviation of measurement error for maturity 30	36	48	60	72	84	96	108	120
KF	7.27 (2.28)	7.91 (2.44)	10.30 (3.00)	9.26 (2.80)	10.04 (3.02)	11.18 (3.37)	10.70 (3.40)	15.07 (4.55)	17.28 (5.12)
PSKF	7.29 (2.29)	7.93 (2.45)	10.30 (3.01)	9.25 (2.80)	10.03 (3.02)	11.14 (3.37)	10.71 (3.40)	15.13 (4.56)	17.29 (5.12)

Table 3: Parameter estimates of decay parameter λ and of standard deviations of measurement errors, expressed in basis points, i.e. $100\sqrt{diag(\Sigma_\epsilon)}$. *KF* denotes the conventional Kalman filter and *PSKF* the pruned skewed Kalman filter. Asymptotic standard errors appear in parenthesis.

DSGE model estimation

DSGE model estimation

The log-linearized model equations of Ireland [2004] are given by:

$$\hat{x}_t = \hat{y}_t - \omega \hat{a}_t$$

$$\hat{g}_t = \hat{y}_t - \hat{y}_{t-1} + \hat{z}_t$$

$$\hat{x}_t = \alpha_x \hat{x}_{t-1} + (1-\alpha_x) E_t \hat{x}_{t+1} - (\hat{r}_t - E_t \hat{\pi}_{t+1}) + (1-\omega)(1-\rho_a)\hat{a}_t$$

$$\hat{\pi}_t = \beta \left(\alpha_\pi \hat{\pi}_{t-1} + (1-\alpha_\pi) E_t \hat{\pi}_{t+1}\right) + \psi \hat{x}_t - \hat{e}_t$$

$$\hat{r}_t - \hat{r}_{t-1} = \rho_\pi \hat{\pi}_t + \rho_x \hat{x}_t + \rho_g \hat{g}_t + \eta_{r,t}$$

$$\hat{a}_t = \rho_a \hat{a}_{t-1} + \eta_{a,t}$$

$$\hat{e}_t = \rho_e \hat{e}_{t-1} + \eta_{e,t}$$

$$\hat{z}_t = \eta_{z,t}$$

where all hat variables are in log deviations from their non-stochastic steady-state.

DSGE model estimation

- Data, same as Ireland [2004], 1980Q1-2003Q1:
  1. Demeaned quarterly changes in seasonally adjusted real GDP per capita
  2. Demeaned quarterly changes in seasonally adjusted real GDP deflator,
  3. Demeaned quarterly averages of daily three-month U.S. Treasury bill rate

- 10 model parameters, 2 are kept fixed

- Estimate the standard deviation and skewness coefficient, directly

- Ensure that means of shocks are zero

- Pruning threshold level: 1%

DSGE model estimation

	Model Parameters			*Shock Parameters*	
	KF	*PSKF*		*KF*	*PSKF*
ω	0.0581 (0.0877)	0.1516 (0.0086)	$skew(\eta_a)$	0	−0.2138 (0.0091)
α_x	0.0000 (0.0043)	0.0002 (0.0004)	$skew(\eta_e)$	0	−0.2227 (0.0198)
α_π	0.0000 (0.0025)	0.0000 (0.0019)	$skew(\eta_z)$	0	−0.9489 (0.0764)
ρ_π	0.3865 (0.1273)	0.2770 (0.0198)	$skew(\eta_r)$	0	0.8028 (0.0437)
ρ_g	0.3960 (0.0650)	0.3402 (0.0130)	$stderr(\eta_a)$	0.0302 (0.0166)	0.0253 (0.0016)
ρ_x	0.1654 (0.0615)	0.2835 (0.0022)	$stderr(\eta_e)$	0.0002 (0.0001)	0.0002 (0.0001)
ρ_a	0.9048 (0.0596)	0.9170 (0.0107)	$stderr(\eta_z)$	0.0089 (0.0015)	0.0079 (0.0003)
ρ_e	0.9907 (0.0155)	0.9812 (0.0250)	$stderr(\eta_r)$	0.0028 (0.0004)	0.0028 (0.0002)
Value of maximized Log-Likelihood function:				1207.56	1215.84

Table 4: Parameter estimates. *KF* denotes the conventional Kalman filter and *PSKF* the pruned skewed Kalman filter. Asymptotic standard errors appear in parenthesis.

- Likelihood ratio test:
  test-statistic = 16.55, p-value = 0.0024
- Hints to problems of linearizing DSGE models

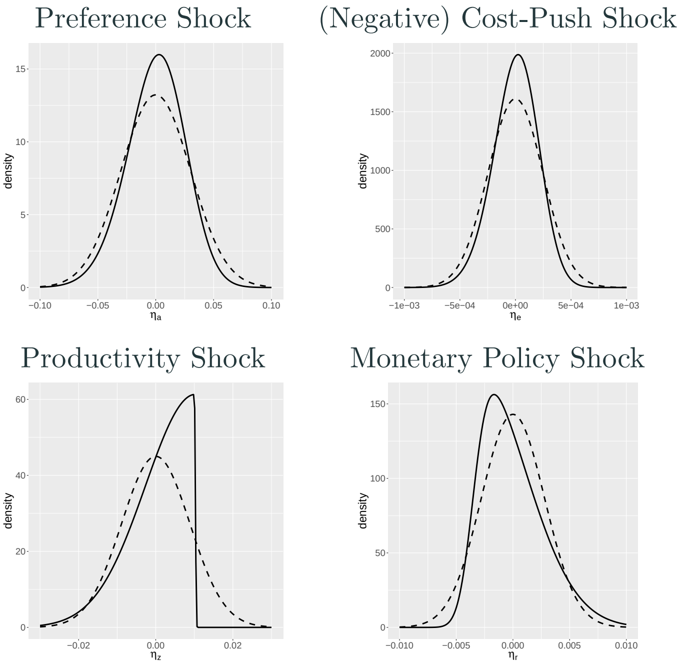


Figure 2: Solid lines are estimated CSN, dashed lines Gaussian

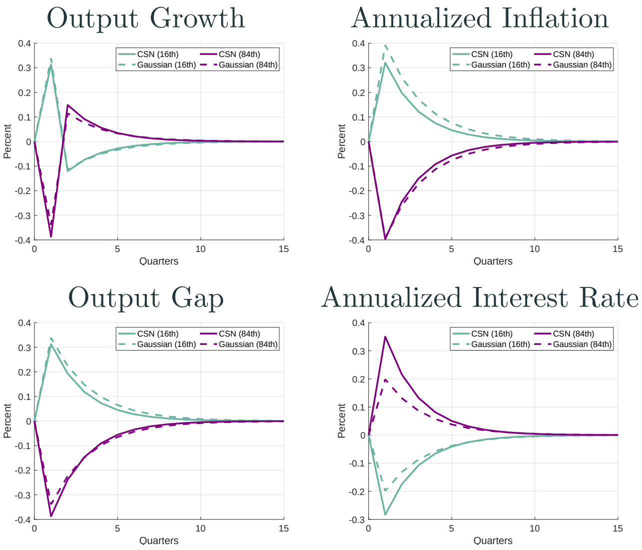


Figure 3: Impulse Responses: Monetary Policy Shock

Conclusion

- Good performance in practice
- Theoretical evidence supporting its validity
- Easy to use
- Compatible with existing toolboxes and standard estimation methods

References

Frank Smets and Rafael Wouters. Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach. *American Economic Review*, 97(3):586–606, 2007. doi: 10.1257/aer.97.3.586.

Ranjith Chiplunkar and Biao Huang. Latent variable modeling and state estimation of non-stationary processes driven by monotonic trends. *Journal of Process Control*, 108:40–54, December 2021. doi: 10.1016/j.jprocont.2021.10.010.

Francis X. Diebold and Canlin Li. Forecasting the term structure of government bond yields. *Journal of Econometrics*, 130(2):337–364, February 2006. doi: 10.1016/j.jeconom.2005.03.005.

Peter N. Ireland. Technology Shocks in the New Keynesian Model. *Review of Economics and Statistics*, 86(4):923–936, November 2004. doi: 10.1162/0034653043125158.