Nonlinear Dynamic Factor Models

Pablo Guerron-Quintana[1,2] · Alexey Khazanov[1] · Molin Zhong[3]

[1]Department of Economics, Boston College
[2]ESPOL
[3]Board of Governors of the Federal Reserve System

Motivation

Linear factor models intensively used in macroeconomics do not capture nonlinearities arising during deep recessions – financial crisis 2008/2009 – or binding constraints – ZLB.

Dynamic Factor Model

We propose a nonlinear dynamic factor model featuring a state equation pruned to the second order based on Andreasen et al. (2013) and a general nonlinear measurement equation.

Implementation and Applications

We use Uscented Kalman Filter and Particle Filter for Maximum Likelihood estimation of the model on US key macroeconomic indicators and cross-country panel of CDS spreads data.

Model

General model:

$$\text{Measurement } y_t = \mathcal{G}(f_t) + \eta\epsilon_t$$
$$\text{Factor motion } f_t = \mathcal{H}(f_{t-1}) + \sigma\nu_t.$$

Here, ϵ_t and ν_t are $iid\ N(0, I)$. Using a generic nonlinear factor motion may lead to

1. Explosive dynamics
2. Divergence of filter

To fix this, we use pruning (Andreasen et al., 2013) – it allows to limit higher order effects. The model is – up to second-order effects:

$$\begin{cases} f_t = f_t^1 + f_t^2 \text{ – 1st and 2nd order factors} \\ f_t^1 = \mathcal{H}_x f_{t-1}^1 + \sigma\nu_t \\ f_t^2 = \mathcal{H}_x f_{t-1}^2 + 0.5\mathcal{H}_{xx}\left(f_{t-1}^1 \times f_{t-1}^1\right). \end{cases} \quad (1)$$

At this point, we don't impose any restrictions on $\mathcal{G}(\cdot)$. Furthermore, our approach can be easily extended to more factors, alternative distributional forms – heteroskedasticity, kurtosis, skewness.

Unscented Kalman Filter – UKF

Approach: Use unscented transform for approximating filtering distributions. UKF forms a Gaussian approximation to the filtering distribution:

$$p(f_t|Y_{1:t}) \approx N(f_t|m_t, P_t).$$

Here, m_t and P_t are mean and covariance. UKF captures first and second moments of the resulting random variables using sigma points.
Algorithm from Särkkä (2013):

- **Prediction**
  1. Given m_{t-1} and P_{t-1}, form sigma points:
  $$\chi_{t-1}^{(0)} = \mathbf{m}_{t-1}, \chi_{t-1}^{(i)} = \mathbf{m}_{t-1} + \sqrt{n+\lambda}\left[\sqrt{\mathbf{P}_{t-1}}\right]_i, \chi_{t-1}^{(i+n)} = \mathbf{m}_{t-1} - \sqrt{n+\lambda}\left[\sqrt{\mathbf{P}_{t-1}}\right]_i,$$
  for $i = 1, \dots, n$,. Here $\lambda = \alpha^2(n+t) - n$ (n – dimensionality of the state), $\mathbf{m}$ is filtered state's (factor's) mean, $\mathbf{P}$ – filtered state's covariance matrix, t denotes period.
  2. Propagate sigma points through the dynamic model:
  $$\hat{\chi}_t^{(i)} = \mathcal{H}(\chi_{t-1}^{(i)}), i = 0, \dots, 2n$$
  3. Compute predicted mean $\mathbf{m}_t^-$ and predicted covariance $\mathbf{P}_t^-$
  $$\mathbf{m}_t^- = \sum_{i=0}^{2n} W_i^{(m)}\hat{\chi}_t^{(i)}, \quad \mathbf{P}_t^- = \sum_{i=0}^{2n} W_i^{(c)}\left(\hat{\chi}_t^{(i)} - \mathbf{m}_t^-\right)\left(\hat{\chi}_t^{(i)} - \mathbf{m}_t^-\right)^T + \mathbf{Q}_{t-1},$$
- **Update step**
  1. Form the sigma points:
  $$\chi_{t-1}^{-(0)} = \mathbf{m}_{t-1}^-, \chi_{t-1}^{-(i)} = \mathbf{m}_{t-1}^- + \sqrt{n+\lambda}\left[\sqrt{\mathbf{P}_{t-1}^-}\right]_i, \chi_{t-1}^{-(i+n)} = \mathbf{m}_{t-1}^- - \sqrt{n+\lambda}\left[\sqrt{\mathbf{P}_{t-1}^-}\right]_i$$
  2. Propagate sigma points through the measurement model:
  $$\hat{y}_t^{(i)} = \mathcal{G}\left(\chi_t^{-(i)}\right), i = 0, \dots, 2n$$
  3. Compute predicted mean μ_t, predicted covariance of the measurement $\mathbf{S}_t$:
  $$\mu_t = \sum_{i=0}^{2n} W_i^{(m)}\hat{\mathcal{Y}}_t^{(i)}, \quad \mathbf{S}_t = \sum_{i=0}^{2n} W_i^{(c)}\left(\hat{\mathcal{Y}}_t^{(i)} - \mu_t\right)\left(\hat{\mathcal{Y}}_t^{(i)} - \mu_t\right)^T + \mathbf{R}_t$$

Advantage: Computationally fast
Disadvantage: Captures only 1st two moments of the distribution
Implementation: Maximum likelihood estimation based on Särkkä (2013) Matlab package

Particle Filter – PF

Approach: generates distribution of particles, weights each particle according to its likelihood, resamples to avoid degeneration and propagates through nonlinear system.
Bootstrap filter version from Särkkä (2013):

- **Prediction**
  Given particles and weights at $t-1$: $\{x_{t-1}^i, W_{t-1}^i\}$
  1. Draw a new particle $x_t^{(i)}$ for each point in the sample set $\{x_{t-1}^{(i)} : i = 1, \dots, N\}$ from
  $$x_t^{(i)} \sim p(x_t|x_{t-1}^{(i)}), \quad i = 1, \dots, N$$
  2. Calculate weights:
  $$\omega_t^{(i)} = p(y_t|x_t^{(i)}), \quad i = 1, \dots, N$$
- **Update**
  1. Define normalized weights: $\widetilde{W}_t^{(i)} = \frac{\omega_t^{(i)}W_{t-1}^{(i)}}{\frac{1}{N}\sum\omega_t^{(i)}W_{t-1}^{(i)}}$.
  2. Resample from multinomial distribution $\left\{\omega_t^{(i)}, \widetilde{W}_t^{(i)}\right\}$ and set $W_t^{(i)} = 1$.

Approximate state distribution and likelihood are:

$$p(x_t|Y_{1:t}) \approx \sum_{i=1}^{N}\omega_t^{(i)}\delta(x_t - x_t^{(i)}), \quad p(y_t|Y_{1:t-1}) \approx \frac{1}{N}\sum_{i=1}^{N}\omega_t^i W_{t-1}^i. \quad (2)$$

Advantage: Tracks the whole distribution
Disadvantage: Computationally and coding-wise heavy
Implementation: Maximum likelihood estimation using CUDA/C++

US Macroeconomic indicators with UKF

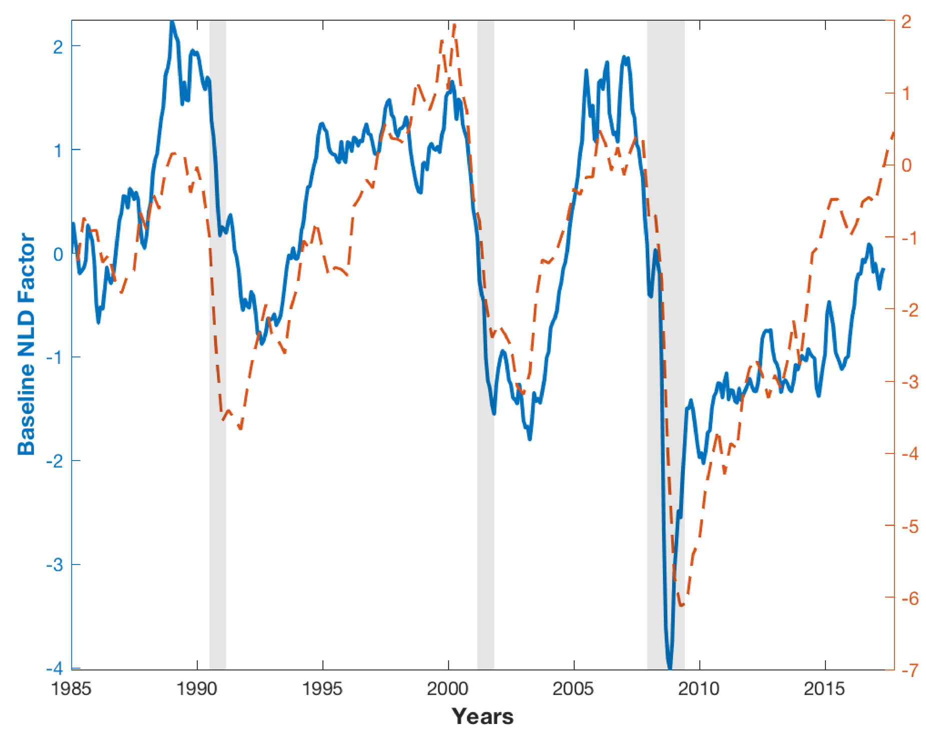
Measurement equation:

$$\begin{bmatrix} R_t \\ y_t \end{bmatrix} = \begin{bmatrix} \max(g_r f_t, -\frac{\mu_r}{\sigma_r}) \\ G \times f_t \end{bmatrix} + \eta\epsilon_t$$

where $\epsilon_t \sim iidN(0, \mathbf{I})$, f follows (1) and $\nu_t \sim iidN(0, 1)$, R is Fed funds rate, y_t includes hourly earnings* , spread between Baa corporate bond yield and 10-year Treasury, CPI inflation*, industrial production index*, spread between 10-Year Treasury Constant Maturity and 2-Year Treasury Constant Maturity, and weekly hours worked, all the series obtained from FRED.

Monthly data covering 1985:1 - 2017:6. Series marked with * were used in log differences $x_t = \ln(\tilde{x}_t) - \ln(\tilde{x}_{t-1})$, where $\tilde{x}$ are original non-stationary series, and all series were standardized: $y_t = \frac{x_t - \bar{x}}{\sigma(x)}$, where x_t are original or log-differentiated series, $\bar{x}$ denotes average and $\sigma(x)$ – standard error. UKF parameters: $n = 2$ (5 sigma points), $\alpha = 1,\ \beta = 0,$ and $\kappa = 1$.

Figure 1. Filtered factor in the estimated model and Output Gap. Data source: FRED, CBO, NBER



Results: The resulting nonlinear economic activity index tracks closely the CBO's output gap.

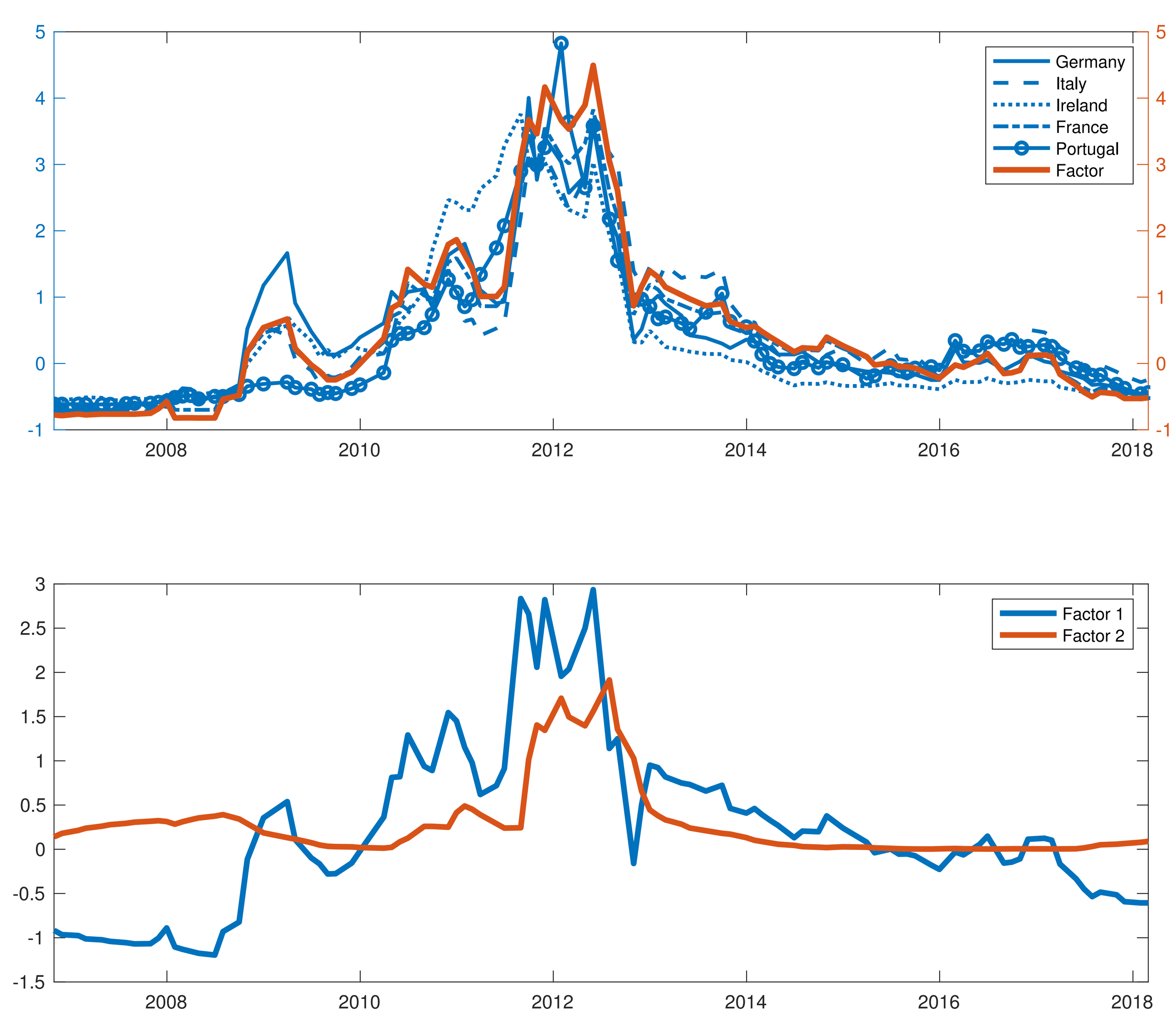
European CDS spreads with PF

Measurement equation:

$$y_t = G \times f_t + \eta\epsilon_t,$$

where y_t are CDS spreads, G is a **5x1** vector, $\epsilon_t \sim iidN(0, \mathbf{I})$ and factor evolves following (1) . We use 5-year USD CDS end of the month spreads obtained from Thomson Reuters Datastream for Germany, Italy, Ireland, France, and Portugal. Sample covers: 2006:10 - 2018:2. All the series were standardized: $y_t = \frac{x_t - \bar{x}}{\sigma(x)}$, where x_t are original series, $\bar{x}$ denotes average and $\sigma(x)$ – standard error. We use $10,000$ particles.

Figure 2. Factor and its decomposition in standardized EU CDS spread series. Data source: Thomson Reuters Datastream



Results: We uncover a common factor that reflects common default risk within the sample of European economies. "Second-order" component of this factor demonstrates significant fluctuations in the period of the sovereign debt crisis, just moderately varying in the rest of the available time interval.

Bibliography

M. M. Andreasen, J. Fernández-Villaverde, and J. Rubio-Ramírez. The pruned state-space system for non-linear dsge models: Theory and empirical applications. Technical report, National Bureau of Economic Research, 2013.

S. Särkkä. *Bayesian filtering and smoothing*, volume 3. Cambridge University Press, 2013.