Dynamic Factor Models & Weekly Economic Trackers

Session 2 · Lecture notes

Tyler Sotomayor

2026-07-30

1 From one indicator to many

Session 1 built a machine — the Kalman filter — that optimally reads *one* noisy indicator of a latent state. Real nowcasting reads dozens. The Federal Reserve's staff track hundreds of series; the FRED-MD research dataset alone curates 127 monthly indicators. This session is about the modeling structure that turns "hundreds of correlated series" from a curse into the entire source of our statistical power.

The starting point is one of the oldest and most robust stylized facts in macroeconomics: **comovement**. When activity turns, it turns *everywhere* — employment, production, income, sales, credit, freight — not identically, but together. Burns and Mitchell (1946) built the NBER's business cycle chronology on this observation decades before anyone could formalize it; Sargent and Sims (1977) supplied the formalization, showing that a small number of common "index" processes account for the bulk of the covariation among U.S. aggregates. The modern statement is blunt: **the business cycle is low-dimensional**. A handful of latent factors — often just one — drive the common movements of everything we can measure.

If that is true, then the right response to "too many series" is not variable selection but *aggregation*: model every series as a noisy witness to a small common state, and let the filter weigh the testimony. You saw the payoff in miniature in the P1 instructor build, where four public series and one factor reproduced the official ten-series Weekly Economic Index with correlation 0.90. This session builds the machinery properly.

2 The static factor model

2.1 The model

Let $x_t \in \mathbb{R}^N$ be a vector of N observed series, standardized, and $f_t \in \mathbb{R}^r$ a vector of $r \ll N$ latent common factors:

$$x_t \;=\; \Lambda f_t + e_t, \tag{1}$$

where Λ $(N \times r)$ holds the **factor loadings** — series i's exposure to each common factor — and e_t is the **idiosyncratic component**: measurement error, sector-specific shocks, everything not common. The core assumptions are that factors and idiosyncratics are uncorrelated, and that the idiosyncratics are (at most weakly) correlated with one another. The *exact* factor model takes $\operatorname{Var}(e_t)$ diagonal; the *approximate* factor model of Bai and Ng (2002) tolerates limited cross-correlation, which is the realistic case — initial and continued claims share labor-market noise no factor should be forced to absorb.

Variance-decompose series i: with everything standardized, $1 = \|\lambda_i\|^2 \operatorname{Var}(f) + \sigma^2_{e,i}$ — each series splits into a **common** share and an **idiosyncratic** share. Good nowcasting indicators are ones with high common shares; a series can be individually fascinating and nearly useless in a factor model because its variation is mostly its own.

2.2 The rotation problem

The factorization in Equation 1 is not unique: for any invertible $r \times r$ matrix M, the pair $(\Lambda M^{-1}, Mf_t)$ produces identical data. The factors are identified only up to rotation, sign, and scale. This is not a defect to be engineered away; it is a fact to be managed by **normalization** — e.g. $\mathrm{Var}(f_t) = I_r$ with $\Lambda'\Lambda$ diagonal, which is what principal components imposes, plus a sign convention (P1: the factor correlates positively with activity). Two honest consequences:

1. **Factors have no intrinsic units.** "The factor fell" means nothing until you scale it against something interpretable — which is exactly why P1 regressed GDP growth on the factor before publishing a number.
2. **Interpretation is a choice.** Calling the first factor "real activity" is justified by its loadings and its behavior, not by the algebra.

2.3 Principal components as estimator

Stack the standardized data in X $(T \times N)$. Principal components estimates $(\hat{\Lambda}, \hat{f}_t)$ by least squares — minimize $\sum_{i,t}(x_{it} - \lambda_i' f_t)^2$ — whose solution is the top r eigenvectors of the sample covariance matrix: exactly the SVD computation in the P1 build. Two classical results make PCA more than a heuristic:

- **Consistency in both dimensions.** As $N, T \to \infty$, the estimated factor space converges to the true one *even under approximate-factor cross-correlation* (Stock and Watson 2002; Bai and Ng 2002). More series genuinely help: each new series is another noisy vote on f_t, and the idiosyncratic noise averages away. This "blessing of dimensionality" is the theoretical license for the everything-plus-the-kitchen-sink datasets of modern nowcasting.
- **Choosing r.** The Bai and Ng (2002) information criteria penalize factor count in a way calibrated to the double-asymptotic setting. In practice, nowcasting models are parsimonious — one or two factors carry U.S. aggregate activity, and the marginal factor tends to pick up sectoral detail (a second factor often loads on housing or finance). The scree plot is not proof, but with the criteria and loadings inspection it is usually decisive.

2.4 What PCA cannot do

PCA is static and complete-data. It has no notion of factor *dynamics* (no forecasting), no native treatment of *missing observations* (no ragged edge, no mixed frequency), and no *uncertainty bands*. All three absences point the same direction: put the factor model in state space and bring back Session 1's filter.

3 The dynamic factor model

3.1 The model

Give the factors dynamics and, where needed, the idiosyncratics too:

$$
\begin{aligned}
x_t &= \Lambda f_t + e_t, \\
f_t &= A_1 f_{t-1} + \cdots + A_p f_{t-p} + u_t, \qquad u_t \sim \mathcal{N}(0, \Sigma_u), \\
e_{it} &= \rho_i \, e_{i,t-1} + \varepsilon_{it}, \qquad \varepsilon_{it} \sim \mathcal{N}(0, \sigma_i^2) \text{ independent across } i.
\end{aligned}
\tag{2}
$$

The factor VAR captures the business cycle's persistence — this quarter's activity predicts next quarter's — and the idiosyncratic AR terms absorb the serial correlation each series carries on its own. Setting $p = 1$, $r = 1$, $\rho_i \neq 0$ is already a serious nowcasting model.

3.2 State space form

Everything from Session 1 applies once we write Equation 2 in state space. For the one-factor, AR(1)-idiosyncratic case, stack the factor *and* the idiosyncratic states:

$$\alpha_t = \begin{pmatrix} f_t \\ e_{1t} \\ \vdots \\ e_{Nt} \end{pmatrix}, \qquad T = \begin{pmatrix} a & & \\ & \rho_1 & \\ & & \ddots \end{pmatrix}, \qquad Z = \begin{pmatrix} \Lambda & I_N \end{pmatrix}, \qquad H = 0.$$

The measurement equation is *exact* ($H = 0$): the "noise" has been promoted into the state, where the filter can model its persistence. The Kalman filter now delivers, in one forward pass: the factor estimate $\mathbb{E}[f_t \mid Y_t]$ with bands, forecasts through the factor VAR, likelihood evaluation — and, because missing observations just delete measurement rows (Session 1, §2.4), the ragged edge is handled *natively*. Add the Mariano–Murasawa tent-weight row for quarterly GDP and the same object is a mixed-frequency nowcasting model: this is, in structure, exactly the model of Giannone et al. (2008).

Dimension check, because it bites in practice: with N series the state has $N + rp$ elements and the filter's cost is cubic in state dimension. For large N, production systems exploit the structure (diagonal blocks, univariate filtering from Session 1 §3.6) or keep only slow-moving series' idiosyncratics in the state.

3.3 Estimation, three ways

(1) Pure principal components. Fast, transparent, no missing data allowed, no dynamics used. The right tool for a first look and for balanced historical panels. This is the P1 build.

(2) Two-step (Doz et al. 2011): estimate Λ and the factor VAR by PCA-plus-OLS on the balanced part of the panel, then *fix those parameters* and run the Kalman filter/smoother over the full ragged panel. One re-filter, no iteration. Remarkably close to fully efficient in realistic designs, and the workhorse for daily production because it is fast and stable.

(3) Maximum likelihood via EM (Doz et al. 2012; Bańbura and Modugno 2014): iterate (E) Kalman smoother for the states given parameters, (M) regressions on smoothed moments for the parameters given states, until the likelihood converges. Handles *arbitrary* missingness patterns in estimation itself, not just in filtering, and yields the full information-matrix machinery. This is the standard behind serious central-bank nowcasting systems. Its cost is implementation care: EM's convergence is monotone but can be slow, and the likelihood surface has the rotation ridge — normalize before you interpret.

The practical sequence in this course mirrors the list: P1 used (1); this session's exercises implement (2) on the same panel; the capstone uses (2) or (3) as ambition allows.

3.4 The news decomposition, now operational

Session 1 stated the news decomposition abstractly. In the DFM it becomes a daily production report. Between vintage Ω_{old} and Ω_{new}, each newly released series i contributes

$$\underbrace{\mathbb{E}[y^* \mid \Omega_{new}] - \mathbb{E}[y^* \mid \Omega_{old}]}_{\text{nowcast revision}} = \sum_{i \in \text{releases}} w_i \underbrace{\left(x_{it} - \mathbb{E}[x_{it} \mid \Omega_{old}]\right)}_{\text{news in release } i},$$

with weights w_i built from the Kalman gain — large for series with high common shares, small publication lags, and precise measurement. This is the sentence "retail sales surprised by +0.4pp and moved the nowcast +0.07pp," derived. Notice the two distinct objects: the *surprise* (data minus model expectation) and the *weight* (how much that series' surprise matters). Commentary that quotes the release against consensus is talking about a different surprise than the model's; a good nowcast writeup is explicit about which expectation was beaten.

4 Anatomy of a weekly tracker

Lewis et al. (2022) is the cleanest published specimen of the whole pipeline, which is why this course replicates it. Its design choices, and why they matter:

Ten weekly series spanning labor (initial and continued claims, staffing index), consumption (Redbook same-store retail sales, consumer confidence), production and shipping (raw steel, fuel sales, electricity output, rail carloads), and taxes (federal withholding). Breadth across *sectors* is the point — the P1 journal's lesson that a claims-only factor makes every claims anomaly a tracker anomaly, formalized.

52-week differences. Weekly data has ferocious and irregular seasonality (holidays float, weeks shift). Differencing against the same week one year back removes it without estimating a seasonal model — at the price of making the index a *trailing-year* growth rate, smoother and slower than the underlying weekly pulse. Session 3 takes up the alternative: actually modeling high-frequency seasonality.

One factor, principal components, GDP scaling. The published index is scaled so a WEI of 2 reads as "GDP growth of about 2% over the trailing year" — the same regression scaling as P1. Refinements in the published version (and good exercises): revisions handling as new weeks arrive, and a COVID-era adjustment after the 2020 outliers distorted the historical scaling — their production version of the standardization trap you have already met.

What it is not. The WEI is a *coincident* index in trailing-year units, not a quarterly-GDP nowcast with a news decomposition; the full DFM of Section 3 is the richer object. The pair — transparent PCA tracker, full-machinery DFM — bracket the design space, and knowing when the simple end suffices is part of the craft.

Cousins worth knowing by name: the Chicago Fed's CFNAI (85 monthly series, PCA — the direct descendant of Stock and Watson (1989)); the ADS index of Aruoba et al. (2009) (mixed daily/weekly/quarterly frequency, exact DFM — closest in spirit to what this course builds); and the generalized DFM literature of Forni et al. (2000), which lets factors load with leads and lags.

5 Practical wisdom

Hard-won rules, several rediscovered live in the P1 build:

1. **Breadth beats depth.** Ten series from ten sectors dominate thirty from three. The factor model averages idiosyncratic noise *only if the noise is actually idiosyncratic* — sectoral duplication correlates it.
2. **The reference population is a modeling choice.** Standardization, loadings, and scaling estimated over a sample containing April 2020 will let one episode define the yardstick (the "standardization trap" — P1 journal, entry 2). Calibrate on normal times, or model the outliers explicitly; never let them in silently.
3. **Transformations do the seasonal work.** 52-week (or 12-month) differences are a seasonal adjustment in disguise. Know what your transformation is *implicitly* assuming before Session 3 makes it explicit.
4. **Inspect loadings like regression coefficients.** A "wrong-signed" or near-zero loading is information: P1's bank-credit series loads 0 because credit-line drawdowns make it acyclical in

year-over-year terms. The model told us something true about the data; listen before deleting.

5. **One factor until proven otherwise.** Extra factors buy sectoral color at the cost of rotation ambiguity and estimation noise. The Bai–Ng criteria are the referee, but parsimony is the prior.

6 Historical note

The intellectual line runs eighty years: Burns and Mitchell (1946)'s reference cycles, distilled by Stock and Watson (1989) into a probabilistic coincident index (their experimental index is a one-factor DFM estimated on four monthly series — recognizably the model in Section 3); scaled to large N by the principal-components consistency results (Stock and Watson 2002; Bai and Ng 2002) and the generalized DFM (Forni et al. 2000); and made operational for the ragged edge by Giannone et al. (2008) and the estimation technology of Doz et al. (2011) and Bańbura and Modugno (2014). The Weekly Economic Index is this whole history compressed into one public time series updated every Thursday morning.

Exercises

1. State space fluency. Write the full system matrices (T, R, Q, Z, H) for a *two*-factor DFM with a VAR(2) factor process, AR(1) idiosyncratics, and $N = 5$ series. What is the state dimension? Where exactly does the Mariano–Murasawa row for quarterly GDP go, and which extra states does it require?

2. Rotation, concretely. Take the P1 panel and its PCA factor. Construct a different loading/factor pair that fits the data identically (any invertible M will do) but whose "factor" would tell a nonsense economic story. State the normalization that rules your construction out.

3. Common shares. For each of the four P1 series, compute the share of variance explained by the factor. Rank the series by usefulness to the tracker. Does the ranking match the loadings? Should it?

4. Two-step upgrade. Implement Doz et al. (2011) on the P1 panel: fix $\hat{\Lambda}$ from PCA, fit an AR(1) to the PCA factor, cast as state space, and run the Kalman smoother. Compare the smoothed factor to the raw PCA factor — where do they differ most, and why does the difference concentrate where data is thin or noisy?

5. A news decomposition by hand. Using your Exercise 4 model: delete the final four weeks of initial claims, filter, record the nowcast; restore one week at a time and recompute. Verify the revisions sum across releases, and that each equals gain × surprise.

6. Choosing r on real data. Download the FRED-MD monthly panel, apply the Bai and Ng (2002) IC_{p2} criterion, and report the chosen number of factors. Plot the loadings of factor 2. Does it have a sectoral interpretation?

7. WEI archaeology (bridge to P1 v2). The published WEI adjusted its methodology during 2020. From Lewis et al. (2022) and the Dallas Fed's documentation, identify what changed and relate it precisely to the standardization trap in the P1 journal. Propose — in one page — the v2 design for our four-series tracker that best imports the lesson.

Further reading

- Stock and Watson (2002) and Bai and Ng (2002) — the two pillars of large-N factor econometrics; read for the theorems' *conditions*, which are where the practice lives.
- Doz et al. (2011) and Doz et al. (2012) — estimation as actually done in production; the two-step paper is the one to implement first.
- Bańbura and Modugno (2014) — EM with arbitrary missingness; the nowcasting workhorse.
- Lewis et al. (2022) — **the practicum paper**; read with the P1 build open.

- Giannone et al. (2008) — reread §3 now that you own the DFM; the model will look familiar in a way it could not have after Session 1.

Aruoba, S. Borağan, Francis X. Diebold, and Chiara Scotti. 2009. "Real-Time Measurement of Business Conditions." *Journal of Business & Economic Statistics* 27 (4): 417–27.

Bai, Jushan, and Serena Ng. 2002. "Determining the Number of Factors in Approximate Factor Models." *Econometrica* 70 (1): 191–221.

Bańbura, Marta, and Michele Modugno. 2014. "Maximum Likelihood Estimation of Factor Models on Datasets with Arbitrary Pattern of Missing Data." *Journal of Applied Econometrics* 29 (1): 133–60.

Burns, Arthur F., and Wesley C. Mitchell. 1946. *Measuring Business Cycles*. National Bureau of Economic Research.

Doz, Catherine, Domenico Giannone, and Lucrezia Reichlin. 2011. "A Two-Step Estimator for Large Approximate Dynamic Factor Models Based on Kalman Filtering." *Journal of Econometrics* 164 (1): 188–205.

Doz, Catherine, Domenico Giannone, and Lucrezia Reichlin. 2012. "A Quasi–Maximum Likelihood Approach for Large, Approximate Dynamic Factor Models." *The Review of Economics and Statistics* 94 (4): 1014–24.

Forni, Mario, Marc Hallin, Marco Lippi, and Lucrezia Reichlin. 2000. "The Generalized Dynamic-Factor Model: Identification and Estimation." *The Review of Economics and Statistics* 82 (4): 540–54.

Giannone, Domenico, Lucrezia Reichlin, and David Small. 2008. "Nowcasting: The Real-Time Informational Content of Macroeconomic Data." *Journal of Monetary Economics* 55 (4): 665–76.

Lewis, Daniel J., Karel Mertens, James H. Stock, and Mihir Trivedi. 2022. "Measuring Real Activity Using a Weekly Economic Index." *Journal of Applied Econometrics* 37 (4): 667–87.

Sargent, Thomas J., and Christopher A. Sims. 1977. "Business Cycle Modeling Without Pretending to Have Too Much a Priori Economic Theory." In *New Methods in Business Cycle Research*. Federal Reserve Bank of Minneapolis.

Stock, James H., and Mark W. Watson. 1989. "New Indexes of Coincident and Leading Economic Indicators." In *NBER Macroeconomics Annual 1989*, edited by Olivier J. Blanchard and Stanley Fischer, vol. 4. MIT Press.

Stock, James H., and Mark W. Watson. 2002. "Forecasting Using Principal Components from a Large Number of Predictors." *Journal of the American Statistical Association* 97 (460): 1167–79.