Are asset pricing models sparse?

Doron Avramov, Guanhao Feng, **Jingyu He** and Shuhua Xiao

Dec 12, 2024
University of Macau

First Macau International Conference on Business Intelligence and Analytics

Sparsity

- Addressing high-dimensional problems is a central challenge in modern statistics.
- Statisticians have developed lots of tools:
  - **Shrinkage:** L_2 penalty.
  - **Selection for sparse models:** L_1 penalty.
- Usually, we assume that the underlying signal is sparse, and advanced methods are designed to recover such signals effectively.
- However, a less frequently explored question arises: Are asset pricing models inherently sparse?

Motivation: Illusion of Sparsity

- Giannone, Lenza, and Primiceri (2021) propose a Bayesian sparse model that parametrizes the level of sparsity.
- They examine various types of economic data, including:
  - **Macro:** Monthly growth rate of U.S. industrial production / GDP.
  - **Finance:** S&P 500 equity premium / stock returns of U.S. firms.
  - **Micro:** Crime rate per capita / the number of pro-plaintiff eminent domain decisions.
- Their findings show that the posterior distribution does not typically concentrate on a single sparse model.
- This phenomenon highlights an illusion of sparsity in economic data.
- They did not emphasize factors.

Evidence from Asset Pricing

The asset pricing literature provides some evidence

- Kozak, Nagel, and Santosh (2020) demonstrate that a characteristics-sparse stochastic discount factor (SDF) cannot explain the cross-section of returns.
- Kozak and Nagel (2023) show that factors derived from characteristics through sorting, characteristic weighting, or OLS cross-sectional regression slopes do not span the stochastic discount factor (SDF) unless a large number of characteristics are used simultaneously.
- Shen and Xiu (2024) prove that when signals are weak, ridge regression outperforms Lasso for prediction.
  - Equivalently, the predictive model might not be sparse.

Research Questions

We investigate sparsity within the framework of the **Characteristics-based Factor Model**:

- Kelly, Pruitt, and Su (2019) introduce observable characteristics as instruments for loadings on latent factors by Instrumented Principal Component Analysis (IPCA).
- We examine whether the results exhibit sparsity in the context of latent factor models.

Our Contribution:

- **Econometric Solution:**
  We propose a novel Bayesian sparse latent conditional factor model.
- **Focus of Analysis:**
  We study the sparsity level of firm characteristics when estimating a conditional latent factor model.

Core Notation: q

Spike-and-slab prior (Mitchell and Beauchamp, 1988; George and McCulloch, 1993), a Bayesian variable selection prior.

$$\begin{cases} P(\beta \neq 0) & = q \\ P(\beta = 0) & = 1 - P(\beta \neq 0) = 1 - q \end{cases}$$

$$\beta = \begin{cases} \mathcal{N}\left(0, \gamma^2\right) \text{ with prob } q & \text{The regressor is chosen.} \\ 0 \text{ with prob } 1 - q & \text{The regressor is not chosen.} \end{cases}$$

Key Current Results

- Under different prior means of q, i.e., $p(\beta_i \neq 0)$, and 4 latent factors, the in-sample posterior mean of q is in the range $[0.65, 0.75]$, i.e., we choose **about 70%** characteristics and their interaction with latent factors in the model.
- Preliminary results. More yet to come.

$$
\begin{aligned}
r_{i,t} &= \boldsymbol{\mu}(\mathbf{z}_{i,t-1}) + \boldsymbol{\beta}(\mathbf{z}_{i,t-1})\mathbf{f}_t + \epsilon_{i,t} \\
\text{where} \quad \boldsymbol{\mu}(\mathbf{z}_{i,t-1}) &= \mu_0 + \boldsymbol{\mu}_1 \mathbf{z}_{i,t-1} \\
\boldsymbol{\beta}(\mathbf{z}_{i,t-1}) &= \boldsymbol{\beta}_0 + \boldsymbol{\beta}_1(\mathbf{I}_K \otimes \mathbf{z}_{i,t-1}), \quad \epsilon_{i,t} \sim \mathcal{N}\left(0, \sigma_i^2\right)
\end{aligned}
\tag{1}
$$

- $r_{i,t}$: return of asset i at time t
- $\mathbf{f}_t$: K latent factors
- $\mathbf{z}_{i,t-1}$: vector, L firm characteristics for asset i at time $t-1$

Plugging the dynamics of $\boldsymbol{\mu}$ and $\boldsymbol{\beta}$ into Model (1):

$$
r_{i,t} = \mu_0 + \boldsymbol{\mu}_1 \mathbf{z}_{i,t-1} + \boldsymbol{\beta}_0 \mathbf{f}_t + \boldsymbol{\beta}_1 [\mathbf{f}_t \otimes \mathbf{z}_{i,t-1}] + \epsilon_{i,t}. \tag{2}
$$

$$r_{i,t} = \mu_0 + \boldsymbol{\mu}_1 \mathbf{z}_{i,t-1} + \boldsymbol{\beta}_0 \mathbf{f}_t + \boldsymbol{\beta}_1 [\mathbf{f}_t \otimes \mathbf{z}_{i,t-1}] + \epsilon_{i,t}.$$

- We assume independent spike-and-slab priors on the regression coefficient Giannone, Lenza, and Primiceri (2021).

$$\boldsymbol{\mu}_1, \boldsymbol{\beta}_1 \overset{iid}{\sim} \begin{cases} \mathcal{N}\left(0, \gamma^2\right) \text{ with prob } q \\ 0 \text{ with prob } 1-q, \end{cases} \qquad \gamma^2 \sim \mathrm{IG}(A/2, B/2)$$

$$\mu_0, \boldsymbol{\beta}_0 \overset{iid}{\sim} \mathcal{N}\left(0, \xi^2\right), \qquad \xi^2 \sim \mathrm{IG}(C/2, D/2)$$

$$q \sim \mathrm{Beta}(a, b), \text{ or, } q = q_{\text{fixed}}$$

where $\boldsymbol{\mu}_1 = [\mu_{1,l}]_{1\le l\le L}$ and $\boldsymbol{\beta}_1 = [\beta_{1,l,k}]_{1\le l\le L, 1\le k\le K}$.

Model comparison via marginal likelihood

- Follow the framework of Barillas and Shanken (2018), Chib, Zeng, and Zhao (2020), comparing different settings by marginal likelihood.
- Marginal likelihood integrates out parameters from the likelihood a priori, addresses parameter uncertainty, and offers regularization of dimension implicitly.
- Different settings for q:
  - i Draw q from the Beta dist. The prior means of q is set to 0.1, 0.5 and 0.9.
  - ii Fixed q at 0.1, 0.5 and 0.9. (Investor perspective)
    Noted that the inverstor believes the sparsity level should be fixed at some values $\neq$ the estimated model would have the same sparisy level.
- Numerical calculate from Gibbs samples, following Chib (1995).

Simulation

Data Generate Process:

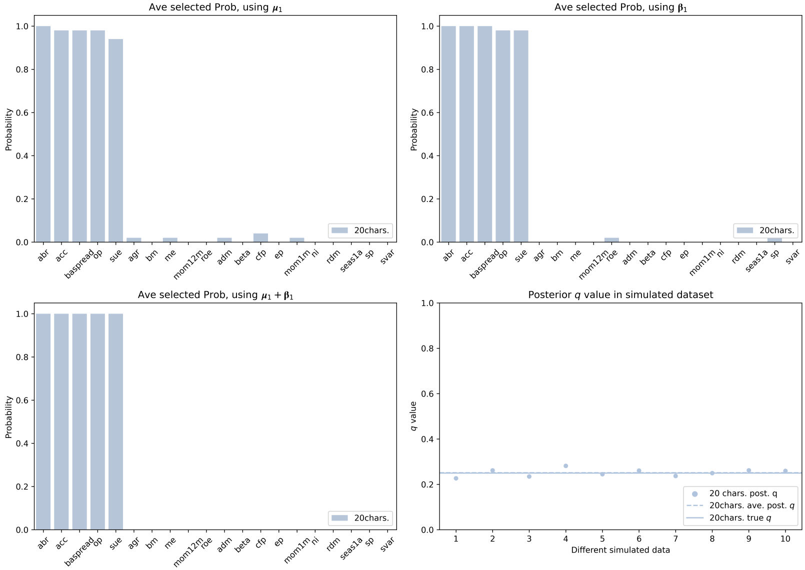
- $\mu_0, \boldsymbol{\mu}_1, \boldsymbol{\beta}_0, \boldsymbol{\beta}_1 \sim \mathcal{N}\left(0, \gamma^2\right)$, where $\gamma^2 \sim \mathcal{IG}\left(20/2, 1/2\right)$
- $\mathbf{f}_t \sim \mathcal{N}\left(0, 0.25^2\right)$
- Generate expected return by using 5 calibrated chars and 3 factors:

$$\mathbb{E}[r_{i,t}] = \mu_0 + \boldsymbol{\mu}_1 \mathbf{z}_{i,t-1} + \boldsymbol{\beta}_0 \mathbf{f}_t + \boldsymbol{\beta}_1 [\mathbf{f}_t \otimes \mathbf{z}_{i,t-1}]$$

- Use Signal-to-Noise Ratio $(= 1)$ to calibrate the return and obtain $r_{i,t}$.

Simulation

Our method can identify the useful (“true”) characteristics.



Current Empirical Findings

- Dataset
- In-sample performance
    - BayesIPCA and IPCA
    - BayesIPCA: Test alpha
    - Sparse BayesIPCA
        - Is there sparsity?
        - Time-varying sparsity

Dataset

360 bi-sorted portfolios, from Jan-1980 to Dec-2023, monthly.

- 1980-2023 monthly observations of U.S. stocks.
- 20 $\mathbf{z}_{i,t}$ firm characteristics (Will be expanded to 60).

BayesIPCA and IPCA

BayesIPCA has a similar pricing performance as IPCA

	Number of factor				
	1	2	3	4	5
Panel A. Total R^2					
IPCA	50.29	69.84	78.43	81.04	82.11
BIPCA	48.94	68.40	77.57	80.20	81.25
Panel B. Pred. R^2					
IPCA	0.39	0.28	0.21	0.24	0.24
BIPCA	0.38	0.24	0.17	0.17	0.20
Panel C. CS R^2					
IPCA	39.86	51.99	57.05	61.14	62.55
BIPCA	41.37	49.04	48.32	56.93	56.89
Panel D. TP. Sp					
IPCA	0.36	0.38	0.79	0.95	0.96
BIPCA	0.38	0.48	0.75	0.98	0.98
Panel E. Uni. Sp					
IPCA	0.36	0.06	0.51	0.14	0.14
BIPCA	0.38	0.27	0.49	0.36	0.15

Benchmark: MktRf.

BayesIPCA: Test alpha

$$r_{i,t} = \underbrace{\mu_0 + \boldsymbol{\mu}_1 \mathbf{z}_{i,t-1}}_{\boldsymbol{\mu}(\mathbf{z}_{i,t-1})} + \underbrace{\beta_0 \mathbf{f}_t + \beta_1 [\mathbf{f}_t \otimes \mathbf{z}_{i,t-1}]}_{\boldsymbol{\beta}(\mathbf{z}_{i,t-1})\mathbf{f}_t} + \epsilon_{i,t}.$$

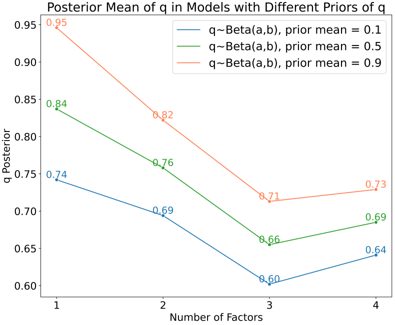
- Test1: Test each μ_0, $\mu_{1,i}, i = 1, \ldots, L$
- Test2: GRS test Gibbons, Ross, and Shanken (1989) on $(\mu_0, \boldsymbol{\mu}_1)$.

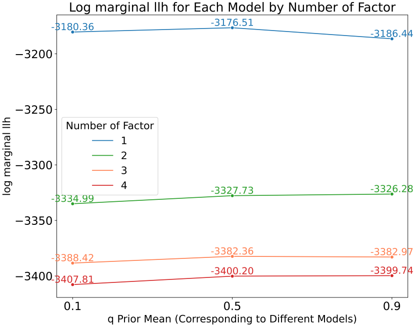
For the in-sample case, K (the number of factors) from 1 to 5, we

- reject the null hypothesis $\mu_i = 0$ in Test1
- reject the null hypothesis $\boldsymbol{\mu} = 0$ in Test2

$\Rightarrow$ There exist some components of returns that cannot be explained by common latent factors and/or characteristics.

Sparse BayesIPCA: Is there sparsity?





Sparse BayesIPCA: Is there sparsity?

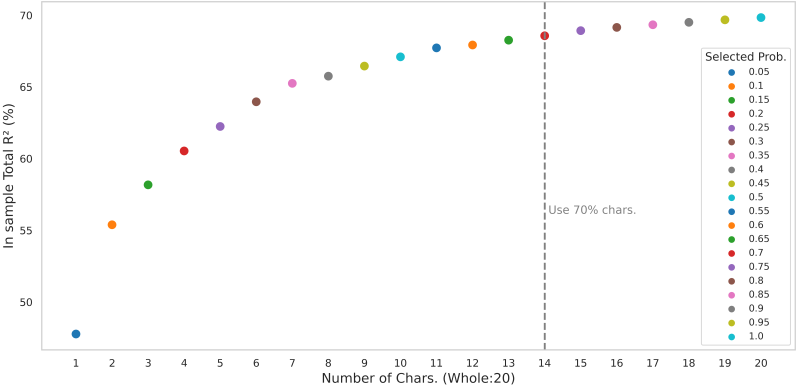
	Number of factor					Number of factor			
	1	2	3	4		1	2	3	4
Panel A. Total R^2									
q prior 0.1	48.94	68.39	77.56	80.16	$q = 0.1$	48.92	68.39	77.55	80.14
q prior 0.5	48.94	68.39	77.56	80.17	$q = 0.5$	48.94	68.39	77.56	80.16
q prior 0.9	48.94	68.39	77.56	80.17	$q = 0.9$	48.94	68.39	77.57	80.18
Panel B. Pred. R^2									
q prior 0.1	0.38	0.22	0.18	0.16	$q = 0.1$	0.38	0.22	0.17	0.15
q prior 0.5	0.38	0.22	0.18	0.16	$q = 0.5$	0.38	0.22	0.17	0.16
q prior 0.9	0.38	0.22	0.18	0.16	$q = 0.9$	0.38	0.22	0.18	0.16
Panel C. CS R^2									
q prior 0.1	41.71	51.40	46.65	53.83	$q = 0.1$	42.46	50.94	46.22	53.95
q prior 0.5	41.59	51.45	46.68	53.89	$q = 0.5$	42.24	51.26	46.54	53.68
q prior 0.9	41.46	51.52	46.66	53.92	$q = 0.9$	41.46	51.59	46.80	54.34
Panel D. TP. Sp									
q prior 0.1	0.39	0.51	0.80	1.03	$q = 0.1$	0.39	0.51	0.79	1.02
q prior 0.5	0.39	0.51	0.81	1.03	$q = 0.5$	0.39	0.51	0.80	1.03
q prior 0.9	0.38	0.51	0.81	1.03	$q = 0.9$	0.39	0.50	0.81	1.02
Panel E. Uni. Sp									
q prior 0.1	0.39	0.32	0.52	0.33	$q = 0.1$	0.39	0.32	0.51	0.32
q prior 0.5	0.39	0.32	0.52	0.33	$q = 0.5$	0.39	0.32	0.51	0.32
q prior 0.9	0.38	0.31	0.52	0.33	$q = 0.9$	0.39	0.30	0.52	0.31

Benchmark: MktRf.

Another Perspective: Construct Sparse Models

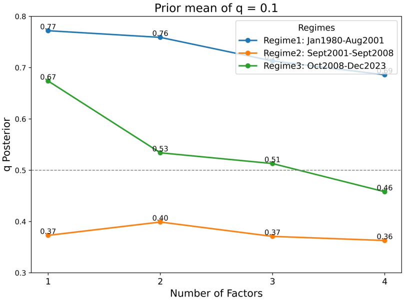
For all possible combinations when select i chars. from 20 chars:

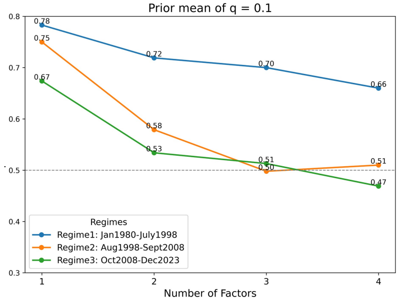
select max(200, C_{20}^{i}) combs., calculate Total R^2, and take the average.



Average in-sample total R^2 (%) by different combs. of chars. (IPCA)

Time-Varying Sparsity





Summary

- An important research problem: **Are the asset pricing models sparse**?
- A new approach, the BayesIPCA Model, combines the **Bayesian framework of factor estimation** and the **characteristics-based model (IPCA)**.
  - An important extension for considering the **spike-and-slab prior** while estimating the conditional latent factor model.
- Based on our method, we can identify:
  - The whole sparsity level of the asset-pricing model (during the whole period / specific regimes)
  - The importance of each characteristic (during the whole period / specific regimes)
  - $\cdots$ The redundancy of the test assets

Thank you!

Technical details

Evaluation Measures

$$\text{Total } R^2 = 1 - \frac{\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left(r_{i,t} - \widehat{r}_{i,t}\right)^2}{\sum_{i=1}^{N}\sum_{t=1}^{T_i}(r_{i,t} - \mathrm{MktRF}_t)^2},$$

where $\widehat{r}_{i,t} = \widehat{\boldsymbol{\mu}}(\mathbf{z}_{i,t-1}) + \widehat{\boldsymbol{\beta}}(\mathbf{z}_{i,t-1})\mathbf{f}_t$.

$$\text{Predictive } R^2 = 1 - \frac{\sum_{i=1}^{N}\sum_{t=1}^{T_i}\left(r_{i,t} - \widehat{r}_{i,t}\right)^2}{\sum_{i=1}^{N}\sum_{t=1}^{T_i}(r_{i,t} - \lambda_{\mathrm{MktRF}})^2},$$

where $\widehat{r}_{i,t} = \widehat{\boldsymbol{\mu}}(\mathbf{z}_{i,t-1}) + \widehat{\boldsymbol{\beta}}(\mathbf{z}_{i,t-1})\boldsymbol{\lambda}_{\mathbf{f}}$, $\boldsymbol{\lambda}_{\mathbf{f}}$ is the factor risk premia estimate, and λ_{MktRF} is the mean of market excess return.

$$\text{Cross-Sectional } R^2 = 1 - \frac{\sum_{i=1}^{N}\left(\frac{1}{T_i}\sum_{t=1}^{T_i}(r_{i,t} - \widehat{r}_{i,t})\right)^2}{\sum_{i=1}^{N}\left(\frac{1}{T_i}\sum_{t=1}^{T_i} r_{i,t} - \mathrm{MktRF}_t\right)^2},$$

where $\widehat{r}_{i,t} = \widehat{\boldsymbol{\mu}}(\mathbf{z}_{i,t-1}) + \widehat{\boldsymbol{\beta}}(\mathbf{z}_{i,t-1})\mathbf{f}_t$.

APT factors

$$f(\mathbf{R} \mid \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}) = \int f^*(\mathbf{R}, \mathbf{f} \mid \boldsymbol{\alpha}, \boldsymbol{\beta}, \boldsymbol{\Sigma}) d\mathbf{f}$$

Gibbs Sampler

The full posterior is

$$
\begin{aligned}
&\text{likelihood} && \prod_{i=1}^{N}\left[\left(\frac{1}{2\pi\sigma_i^2}\right)^{\frac{T_i}{2}}\exp\left(-\frac{1}{2\sigma_i^2}\left(R_i-\mathcal{W}_i\Gamma\right)^{\top}\left(R_i-\mathcal{W}_i\Gamma\right)\right)\right]\\
&\text{prior on } \boldsymbol{\mu}_1 && \times\prod_{l=1}^{L}\left[\left(\frac{1}{2\pi\gamma^2}\right)^{\frac{1}{2}}\exp\left(-\frac{\mu_{1,l}^2}{2\gamma^2}\right)\right]^{z_l^{\mu}}\left[\delta(\mu_{1,l})\right]^{1-z_l^{\mu}}\\
&\text{prior on } \boldsymbol{\beta}_1 && \times\prod_{l=1}^{L}\prod_{k=1}^{K}\left[\left(\frac{1}{2\pi\gamma^2}\right)^{\frac{1}{2}}\exp\left(-\frac{\beta_{1,l,k}^2}{2\gamma^2}\right)\right]^{z_{l,k}^{\beta}}\left[\delta(\beta_{1,l,k})\right]^{1-z_{l,k}^{\beta}}\\
&\text{prior on } \mu_0, \boldsymbol{\beta}_0 && \times\left[\left(\frac{1}{2\pi\xi^2}\right)^{\frac{1}{2}}\exp\left(-\frac{\mu_0^2}{2\xi^2}\right)\right]\times\prod_{l=1}^{L}\left[\left(\frac{1}{2\pi\xi^2}\right)^{\frac{1}{2}}\exp\left(-\frac{\beta_{0,l}^2}{2\xi^2}\right)\right]\\
&\text{prior on } z^{\mu}, z^{\beta}, q && \times\left[\prod_{l=1}^{L}q^{z_l^{\mu}}(1-q)^{1-z_l^{\mu}}\right]\times\left[\prod_{l=1}^{L}\prod_{k=1}^{K}q^{z_{l,k}^{\beta}}(1-q)^{1-z_{l,k}^{\beta}}\right]\times\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}q^{a-1}(1-q)^{b-1}\\
&\text{prior on } \sigma_i^2, \gamma^2 && \times\prod_{i=1}^{N}(\sigma_i^2)^{-\frac{v_0}{2}-1}\exp\left(-\frac{S_0}{2\sigma_i^2}\right)\times\frac{(B/2)^{A/2}}{\Gamma(A/2)}(\gamma^2)^{-A/2-1}\exp\left(-\frac{B}{2\gamma^2}\right)\\
&\text{prior on } \xi^2 && \times\frac{(D/2)^{C/2}}{\Gamma(C/2)}(\xi^2)^{-C/2-1}\exp\left(-\frac{D}{2\xi^2}\right)
\end{aligned}
$$

Gibbs Sampler

For BayesIPCA-sparsity case:

- Sample $p(\tilde{\Gamma} \mid z, \sigma_i^2, \gamma^2, \xi^2)$
- Sample $p(\sigma_i^2 \mid z, \tilde{\Gamma})$
- Sample $p(z \mid \sigma_i^2, \gamma^2, \xi^2, q)$
- Sample $p(\gamma^2 \mid z, \tilde{\boldsymbol{\mu}}_1, \tilde{\boldsymbol{\beta}}_1)$
- Sample $p(q \mid z)$
- Sample $p(\xi^2 \mid \mu_0, \boldsymbol{\beta}_0)$

where $\tilde{\Gamma} = \left(\mu_0, \tilde{\boldsymbol{\mu}}_1, \boldsymbol{\beta}_0, \tilde{\boldsymbol{\beta}}_1\right)$, $\tilde{}$ means the selected variables.

Review: Bayesian APT (Arbitrage Pricing Theory) Factor Model

Geweke and Zhou (1996)

$$\mathbf{r}_t = \boldsymbol{\mu} + \boldsymbol{\beta}\mathbf{f}_t + \boldsymbol{\epsilon}_t$$

- $\mathbf{r}_t = (r_{1,t}, \cdots, r_{N,t})$: a vector of returns of N asset at time t
- $\boldsymbol{\mu} = \mathbb{E}[\mathbf{r}_t]$, the expected return on asset.
- "pervasive" factor assumptions:

  $\mathbb{E}[\mathbf{f}_t] = 0$, $\mathbb{E}[\mathbf{f}_t\mathbf{f}_t'] = \mathbf{I}$, $\mathbb{E}(\boldsymbol{\epsilon}_t \mid \mathbf{f}_t) = 0$, $\mathbb{E}[\boldsymbol{\epsilon}_t\boldsymbol{\epsilon}_t' \mid \mathbf{f}_t] = \boldsymbol{\Sigma}$.
- Gibb sampler, draw $\boldsymbol{\mu}$, $\boldsymbol{\beta}$ and $\boldsymbol{\Sigma}$.
- $\mathbf{f}_t$ and $\mathbf{r}_t$ are jointly normally distributed.

  Draw $\mathbf{f}$ conditional on $\boldsymbol{\mu}$, $\boldsymbol{\beta}$, $\boldsymbol{\Sigma}$ and the data:

$$\begin{pmatrix} \mathbf{f}_t \\ \mathbf{r}_t \end{pmatrix} \sim \mathcal{N}\left[\begin{pmatrix} \mathbf{0} \\ \boldsymbol{\mu} \end{pmatrix}, \begin{pmatrix} \mathbf{I} & \boldsymbol{\beta}' \\ \boldsymbol{\beta} & \boldsymbol{\beta}\boldsymbol{\beta}' + \boldsymbol{\Sigma} \end{pmatrix} \right].$$

$$\mathbb{E}(\mathbf{f}_t \mid \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \mathbf{r}_t) = \boldsymbol{\beta}'(\boldsymbol{\beta}\boldsymbol{\beta}' + \boldsymbol{\Sigma})^{-1}(\mathbf{r}_t - \boldsymbol{\mu}),$$

$$\mathrm{Cov}(\mathbf{f}_t \mid \boldsymbol{\mu}, \boldsymbol{\beta}, \boldsymbol{\Sigma}, \mathbf{r}_t) = \mathbf{I} - \boldsymbol{\beta}'(\boldsymbol{\beta}\boldsymbol{\beta}' + \boldsymbol{\Sigma})^{-1}\boldsymbol{\beta}.$$

Review: IPCA

Kelly, Pruitt, and Su (2019)

$$r_{i,t} = \mathbf{z}'_{i,t-1}\Gamma_\alpha + \mathbf{z}'_{i,t-1}\Gamma_\beta \mathbf{f}_t + \epsilon_{i,t}$$

$$r_{i,t} = \boldsymbol{\mu}(\mathbf{z}_{i,t-1}) + \boldsymbol{\beta}(\mathbf{z}_{i,t-1})\mathbf{f}_t + \epsilon_{i,t}$$

where $\boldsymbol{\mu}(\mathbf{z}_{i,t-1}) = \mathbf{z}'_{i,t-1}\Gamma_\alpha = \boldsymbol{\mu}_1 \mathbf{z}_{i,t-1}$

$$\boldsymbol{\beta}(\mathbf{z}_{i,t-1}) = \mathbf{z}'_{i,t-1}\Gamma_\beta = \boldsymbol{\beta}_1(\mathbf{I}_K \otimes \mathbf{z}_{i,t-1})$$

Review: IPCA

Kelly, Pruitt, and Su (2019)

$$r_{i,t} = \mathbf{z}'_{i,t-1}\Gamma_\alpha + \mathbf{z}'_{i,t-1}\Gamma_\beta \mathbf{f}_t + \epsilon_{i,t}$$

$$r_{i,t} = \boldsymbol{\mu}(\mathbf{z}_{i,t-1}) + \boldsymbol{\beta}(\mathbf{z}_{i,t-1})\mathbf{f}_t + \epsilon_{i,t}$$

$$\text{where} \quad \boldsymbol{\mu}(\mathbf{z}_{i,t-1}) = \mathbf{z}'_{i,t-1}\Gamma_\alpha = \boldsymbol{\mu}_1 \mathbf{z}_{i,t-1}$$

$$\boldsymbol{\beta}(\mathbf{z}_{i,t-1}) = \mathbf{z}'_{i,t-1}\Gamma_\beta = \boldsymbol{\beta}_1(\mathbf{I}_K \otimes \mathbf{z}_{i,t-1})$$

- Estimate of $\boldsymbol{\mu}_1$, $\boldsymbol{\beta}_1$ and $\mathbf{f}_t$ by optimization:

$$\min_{\Gamma_\beta, \Gamma_\alpha, F} \sum_{t=1}^{T} \left(r_t - Z_{t-1}\Gamma_\beta f_t - Z_{t-1}\Gamma_\alpha\right)' \left(r_t - Z_{t-1}\Gamma_\beta f_t - Z_{t-1}\Gamma_\alpha\right).$$

- Method: Alternating Least Square (ALS)
- Some conclusions:
  - Dynamic betas (parameterized functions of observable characteristics)
  - Accept $\boldsymbol{\mu}_1 = \mathbf{0}$.

Bi-sorted portfolio construction

(1) Split the stocks into two groups based on lag me, (the smallest 70% in one group and the largest 30% in the other). Then, further divide each group into three categories based on the ranked standardized characteristics, specifically within the intervals of -1 to -0.4, -0.4 to 0.4, and 0.4 to 1.

(2) Apply value weighting within each decile to obtain weight char. and returns. $\Rightarrow 2 \times 3 \times 10 = 360$

(3) Standardize the characteristics in the cross-section into Uniform$[-1, 1]$.

References

Barillas, F. and J. Shanken (2018). Comparing asset pricing models. *Journal of Finance 73*(2), 715–754.

Chib, S. (1995). Marginal likelihood from the gibbs output. *Journal of the American Statistical Association 90*(432), 1313–1321.

Chib, S., X. Zeng, and L. Zhao (2020). On comparing asset pricing models. *Journal of Finance 75*(1), 551–577.

George, E. I. and R. E. McCulloch (1993). Variable selection via gibbs sampling. *Journal of the American Statistical Association 88*(423), 881–889.

Geweke, J. and G. Zhou (1996). Measuring the pricing error of the arbitrage pricing theory. *Review of Financial Studies 9*(2), 557–587.

Giannone, D., M. Lenza, and G. E. Primiceri (2021). Economic predictions with big data: The illusion of sparsity. *Econometrica 89*(5), 2409–2437.

References ii

Gibbons, M. R., S. A. Ross, and J. Shanken (1989). A test of the efficiency of a given portfolio. *Econometrica*, 1121–1152.

Kelly, B. T., S. Pruitt, and Y. Su (2019). Characteristics are covariances: A unified model of risk and return. *Journal of Financial Economics 134*(3), 501–524.

Kozak, S. and S. Nagel (2023). When do cross-sectional asset pricing factors span the stochastic discount factor? Technical report, National Bureau of Economic Research.

Kozak, S., S. Nagel, and S. Santosh (2020). Shrinking the cross-section. *Journal of Financial Economics 135*(2), 271–292.

Mitchell, T. J. and J. J. Beauchamp (1988). Bayesian variable selection in linear regression. *Journal of the American Statistical Association 83*(404), 1023–1032.

Shen, Z. and D. Xiu (2024). Can machines learn weak signals? *University of Chicago, Becker Friedman Institute for Economics Working Paper* (2024-29).