B. The SDF and the Tangent Portfolio

The SDF can be linked to the mean-variance efficient tangent portfolio. As [3] shows,

$$M_{t+1} = 1 - F_{t+1},$$

where

$$F_{t+1} = \frac{(\mathbf{w}^T)'\mathbf{r}^e_{t+1}}{\gamma_t},$$

for vector of excess returns $\mathbf{r}^e_t$. That is, F_{t+1} is the return on the mean-variance efficient tangent portfolio scaled by a time-varying coefficient $\frac{1}{\gamma_t}$. As a result,

$$M_{t+1} = 1 - \frac{(\mathbf{w}^T)'\mathbf{r}^e_{t+1}}{\gamma_t}. \tag{2}$$

C. Estimating the SDF

In practice, the SDF can be estimated simply by using the Equation (1). Historical returns are known, so the SDF can be fit by the problem

$$\min_{\boldsymbol{\theta}} \sum_i \sum_t (M_{t+1}(\boldsymbol{\theta}) r^e_{i,t+1})^2,$$

minimizing the squared mispricing error. This formulation can be made conditional on inputs $\mathbf{x}$ (proxies for ω), making the problem

$$\min_{\boldsymbol{\theta}} \sum_i \sum_t (M_{t+1}(\mathbf{x};\boldsymbol{\theta}) r^e_{i,t+1})^2.$$

This empirical estimation can be modified into an asset management strategy by using the relation in Equation (2). If one wants to invest in the tangent portfolio, then one can learn a model for its weights by defining

$$M_{t+1}(\mathbf{x};\boldsymbol{\theta},\gamma) = 1 - \frac{(\mathbf{w}^T(\mathbf{x};\boldsymbol{\theta}))'\mathbf{r}^e_{t+1}}{\gamma}.$$

The problem is now

$$\min_{\boldsymbol{\theta},\gamma} \sum_i \sum_t (1 - \frac{(\mathbf{w}^T(\mathbf{x};\boldsymbol{\theta}))'\mathbf{r}^e_{t+1}}{\gamma})^2 (r^e_{i,t+1})^2.$$

IV. Technical Implementation

To improve model robustness, we introduce an L2 regularization term into our objective function. We weight this term against the objective with a constant coefficient λ. Our final implementation uses this combined objective, addressing the problem

$$\min_{\boldsymbol{\theta},\gamma} \sum_i \sum_t (1 - \frac{(\mathbf{w}^T(\mathbf{x};\boldsymbol{\theta}))'\mathbf{r}^e_{t+1}}{\gamma})^2 (r^e_{i,t+1})^2 + \frac{\lambda}{10}||\boldsymbol{\theta}||^2. \tag{3}$$

Using the dataset provided in class for 65 real estate companies over 97 months from 2015 to 2023, we jointly estimate a model of the SDF and tangent portfolio according to Equation (3). Our dataset includes 3 fundamental variables for each company and 3 macroeconomic factors in addition to each company's excess returns (descriptive statistics are shown in the Appendix in Table II). We normalize each factor before inputting them to the model. We train on values between 2015 until the beginning of 2021. We validate the model using values between the beginning of 2021 and the beginning of 2023.

Our implementation uses gradient descent (in the style of Algorithm 1) to minimize the objective. First, we define factor loadings $\mathbf{b}$ and scaling factor γ, both initialized randomly. We use these parameters to calculate

$$\boldsymbol{\omega}_t(\mathbf{X};\mathbf{b}) = \mathbf{X_t b}$$

and

$$M_{t+1}(\mathbf{X};\mathbf{b}) = 1 - \frac{\boldsymbol{\omega}_t'(\mathbf{X};\mathbf{b})\mathbf{r}^e_{t+1}}{\gamma},$$

where $\mathbf{X}_{t,i,k}$ is a three-dimensional tensor containing each value of factor k for each company i at time t and $\boldsymbol{\omega}_t$ is an approximator of $\mathbf{w}^T$ and should not be confused with random state ω discussed above. We use the Adam optimizer to learn these parameters, modifying the parameter update step in Algorithm 1.

Algorithm 1 Gradient Descent Procedure

Input:
- observations $\mathbf{X}$
- iterations N
- regularization coefficient λ
- learning rate η

Output: Estimated parameters $\mathbf{b}^{(N)}$

Initialize $\mathbf{b}^{(0)} \sim \mathcal{U}(0,1)^6$
Define $\mathcal{L}(M_{t+1},\mathbf{b}) = \sum_i \sum_t (M_{t+1})^2 (r^e_{i,t+1})^2 + \frac{\lambda}{10}||\mathbf{b}||^2$
Initialize $\gamma^{(0)} \leftarrow 1$
for $n = 0$ **to** N **do**
  $\boldsymbol{\omega}_t^{(n)} \leftarrow \mathbf{X_t b}^{(n)}$
  $M_{t+1}^{(n)} \leftarrow 1 - \frac{(\boldsymbol{\omega}_t^{(n)})'\mathbf{r}^e_{t+1}}{\gamma^{(n)}}$
  $g_\gamma^{(n)}, g_\mathbf{b}^{(n)} \leftarrow \nabla_{\gamma,\mathbf{b}} \mathcal{L}(M_{t+1}^{(n)},\mathbf{b})$
  $\gamma^{(n+1)} \leftarrow \gamma^{(n)} - \eta g_\gamma^{(n)}$
  $\mathbf{b}^{(n+1)} \leftarrow \mathbf{b}^{(n)} - \eta g_\mathbf{b}^{(n)}$
end for
return $\mathbf{b}^{(T)}$

V. Results

We test our optimization using several different gradient descent algorithms (training losses shown in Figure 1 and validation results shown in Figure 2). We find that Adam is the most effective optimizer, converging to a better solution than stochastic gradient descent (although it does so more slowly). This best model achieves a total squared mispricing error of 2.525×10^3.

To verify our model, we observe the predicted stochastic discount factor over our validation set (Figure 7 in the Appendix). We find that M_{t+1} takes positive values and fluctuates significantly, indicating that our model is outputting a valid SDF that varies with our data. Our estimates of factor loadings can be found in Table I and are shown in Figure 3. The most significant loading is on Extraordinary Items and Discontinued Operations with New House Building Permits