QUADRATIC HEDGING STRATEGIES

For private equity funds

Christian Tausch

April 25, 2021

AssetMetrix

Agenda

1. Introduction
2. Review: Asset pricing frameworks in incomplete markets
3. Model: A quadratic hedging approach for PEF cash flows
4. Application: Componetwise L2 boosting
5. Conclusion

ABSTRACT

In this talk, we first summarize the most common asset pricing frameworks in incomplete markets: using the growth-optimal portfolio as numeraire, mean-variance hedging, and local risk minimization. On their basis, we develop a tailored quadratic hedging framework for private equity fund cash flows. Finally, we demonstrate how to empirically estimate these hedging strategies via the machine-learning method componentwise L2 boosting.

1 INTRODUCTION

Stochastic Discount Factor Methods for Non-Traded Cash Flows - The Case of Private Equity

1. Can we replicate private equity cash flows by ”liquid alternatives”?
2. Compare SDF approaches:

Econometrics: In the semiparametric SDF paper, we incorporate factor returns into the SDF factor construction $\Psi_{\mathrm{PE}}(\mathsf{F})$.

$$\mathbb{E}\left[\sum \Psi_{\mathrm{PE}}(\mathsf{F}) \cdot \mathsf{CF}\right] = 0$$

Financial Mathematics: Now, we use a universal SDF proxy without factors, but use the factors to form a dynamic hedging strategy G(F).

$$\mathbb{E}\left[\sum \frac{\mathsf{CF} - \mathsf{G}(\mathsf{F})}{\frac{1}{\Psi_{\mathrm{universal}}}}\right] = 0$$

Which approach is more intuitive, theoretically more valid, easier to apply?

Let $(\Omega, \mathcal{F}, \mathbb{P})$ be a discrete-time probability space, $T \in \mathbb{N} > 0$, and $\mathbb{F} = (\mathcal{F}_t)_{t=0,1,\ldots,T}$ a filtration.

In our security market $\{S_{z,t} : z = 1, 2, \ldots, Z \;\; ; \;\; t = 0, 1, 2, \ldots, T\}$, the underlying securities S_z are assumed to be real-valued, $\mathbb{F}$-adapted, and square-integrable processes with $S_{z,t} \in [0, \infty)$ for all z, t.

Additionally, one asset is assumed to be default-free (risk-free) $S_{1,t} > 0$ for all t.

We regard portfolios

$$S_t^\delta = {\delta_t}^\top S_t$$

that are all **non-negative**, **self-financing** and **finite**, i.e., $S_t^\delta \in [0, \infty)$ for all t. The strategy vector is denoted by $\delta_t \in \mathbb{R}^N$, δ_t is $\mathcal{F}_t$-measurable.

From the underlying securities S, we can construct zero-net-investment factor returns

$$F_{j,t} = \frac{S_t^{\delta,j,\mathrm{long}}}{S_{t-1}^{\delta,j,\mathrm{long}}} - \frac{S_t^{\delta,j,\mathrm{short}}}{S_{t-1}^{\delta,j,\mathrm{short}}}$$

with $j = 1, 2, ..., J$.

Additionally, relevant public macro information is stored in so-called predictor variables $P_{k,t}$ with $k = 1, 2, \ldots, K$.

Private equity funds are described by their cash flows $CF_{i,t}$ and net asset values $V_{i,t}$ with $i = 1, 2, \ldots, N$ and $t = 1, 2, \ldots, T$.

2 PRICING IN INCOMPLETE MARKETS

Incomplete markets: non-traded assets, trading frictions, asset prices with jumps, ...

We all know: When hedging in incomplete markets one has to sacrifice (a) perfect replication or (b) the self-financing property (which jointly only work in complete markets).

Financial mathematics concepts for pricing in incomplete markets:

1. Numeraire denomination by Growth Optimal Portfolio (GOP)

   –> [Kaplan and Schoar, 2005] PME

2. Self-financing portfolio replication by Mean-Variance Hedging

   –> modified PME (mPME) of Cambridge Associates

3. Perfect portfolio replication by (Local) Risk Minimization

   –> [Tausch, 2019] Quadratic Hedging & PME+ by Capital Dynamics

Definition

Growth-Optimal Portfolio (GOP):
The GOP $S_t^{\delta^*}$ is the unique **strictly positive** portfolio that makes every GOP-denoted portfolio process $\hat{S}_t^\delta = S_t^\delta (S_t^{\delta^*})^{-1}$ a **supermartingale** under the real-world probability measure $\mathbb{P}$ [Bühlmann and Platen, 2003, equation 4.2]:

$$\hat{S}_t^\delta = \frac{S_t^\delta}{S_t^{\delta^*}} \geq \mathbb{E}^{\mathbb{P}} \left[\hat{S}_{t+1}^\delta | \mathcal{F}_t \right] \quad \forall \quad t, \delta \tag{1}$$

The GOP framework requires only one trivial assumption:
Assumption: The GOP exits [Bühlmann and Platen, 2003, assumption 3.2].

Helpful additional assumption: every GOP-denominated portfolio process is a **martingale**.

Mean-variance hedging in discrete-time can be described by the following **global** minimization task [Schweizer, 1995]

$$\min_{c\in\mathbb{R},\vartheta\in\Theta} \mathbb{E}^{\mathbb{P}}\left[(V_T - c - G_T(\vartheta))^2\right] \tag{2}$$

where Θ includes the set of all "admissible" **self-financing** trading strategies and the cumulative gain process $G(\vartheta)$ is defined as

$$G_t(\vartheta) = \sum_{\tau=1}^{t} {\vartheta_{\tau-1}}^{\top}(\hat{S}_\tau - \hat{S}_{\tau-1}) \tag{3}$$

where we use GOP-denominated asset price process $\hat{S}_t = \frac{S_t}{S_t^{\delta*}}$ to operate in a convenient martingale setting.

- Local risk minimization aims at **perfect replication** of a given payoff by means of a **mean-self-financing strategy**
- Expected cost of trading strategy is zero for all t
- Think of it as ”hedging by sequential regression”
- $V_t = (\varphi_{t-1})^\top \hat{S}_t + \eta_t$
- Trading cost (= hedging error): $\mathbb{E}^{\mathbb{P}}[\eta_t] = 0$ for all t
- Local risk minimization: discounted S is a semimartingale (= martingale + adapted finite-variation process)
- Risk minimization: discounted S is a martingale
- Literature for (local) risk minimization of a payment stream process [Moller, 2001, Pansera, 2012]

3 QUADRATIC HEDGING FOR PEFS

We want to estimate a dynamic hedging strategy $\varphi \in \mathbb{R}^{T \times Z}$ tailored for private equity fund cash flows!

1. As $T \times N$ can be large, use high-dimensional statistics
2. To reduce N, use factor returns (F) instead of single stocks (S)
3. To reduce T, use dynamic predictors (P)
4. To avoid change of measure, discount by GOP $S^{\delta *} = \frac{1}{\Psi_{\mathrm{GOP}}}$
5. To get a tractable optimization problem, consider mean-self-financing (LRM) instead of self-financing (MV) hedges
6. To not rely too much on NAV, use global loss criterion (at time T_i)
7. To start with zero initial investment, use zero-net-investment factor returns (F)

$$\mathbb{E}\left[\sum \frac{\mathrm{CF} - G_{\mathrm{RM}}(F, P)}{\frac{1}{\Psi_{\mathrm{GOP}}}}\right] = 0$$

Given observed fund NAV $\tilde{V}$ and cash flows $\tilde{CF}$ we define

$$V_{i,t} := \frac{\tilde{V}_{i,t}}{S_t^{\delta *}}$$

and the corresponding cumulative cash flow

$$A_{i,t} := \sum_{\tau=1}^{t} \frac{\tilde{CF}_{i,\tau}}{S_\tau^{\delta *}}$$

The cumulative gain function associated with a given hedging strategy is given by

$$G_t := \sum_{\tau=1}^{t} \sum_{j=1}^{J} \xi_{j,\tau} \frac{F_{j,\tau} - \lambda}{S_\tau^{\delta *}} \tag{4}$$

relying on the linear hedging strategy ξ

$$\xi_{j,\tau} = \sum_{k=1}^{K} \beta_{j,k} \cdot P_{\tau-1}^{(k)} \cdot \tilde{V}_{\tau-1}$$

with trading cost λ, coefficients β, predictors P, factor returns F.

Next, the replication target of our hedging strategy is simply

$$Y_{i,t} = V_{i,t} + A_{i,t} \tag{5}$$

We favor the squared final hedging error loss function

$$L_i = \left(Y_{i,T_i} - G_{i,T_i}\right)^2 \tag{6}$$

instead of the risk minimizing loss function that averages over all t.

We determine the optimal β coefficient vector based on the empirical loss function estimate

$$\beta^* = \arg\min_{\beta} \frac{1}{N} \sum_{i=1}^{N} L_i\left(\beta\right)$$

with $\beta \in \mathbb{R}^{K \times J}$ which shall be easier to estimate than $\varphi \in \mathbb{R}^{T \times Z}$.

4 COMPONENTWISE L2 BOOSTING

4.1 (COMPONENTWISE) BOOSTING IDEA

- Boosting is a family of machine learning algorithms that convert weak learners to strong ones
- Boosting is an iterative procedure where in each step the residual error is further minimized (”regression on error term”)
- In our case, the residual error is the **residual hedging error**
- In componentwise boosting in each step, we select the ”component” that reduces the residual error the most
- In our case, we have $J \times K$ components as each factor-predictor pair is considered a component
- componentwise L^2 boosting uses a quadratic loss function
- Boosting too many iteration yields a too strong model with a too small (in-sample) error, you must stop early to not overfit

Select the univariate factor-predictor combination with maximal explanatory power

$$\hat{j}, \hat{k} = \mathop{\text{arg min}}_{j,k} \frac{1}{N} \sum_{i=1}^{N} \left(u_i - \hat{g}_i^{(j,k)} \right)^2 \tag{7}$$

with pseudo-response variable u_i and optimal univariate gain function analogue to equation (4)

$$\hat{g}_i^{(j,k)} = \sum_{\tau=1}^{T_i} \hat{\beta}_{j,k} \cdot P_{\tau-1}^{(k)} \cdot \tilde{V}_{i,\tau-1} \cdot \frac{F_{j,\tau} - \lambda}{S_\tau^*} \tag{8}$$

As prerequisite, estimate the optimal univariate β coefficient for each given factor-predictor combination j, k

$$\hat{\beta}_{j,k} = \mathop{\text{arg min}}_{\beta_{j,k}} \frac{1}{N} \sum_{i=1}^{N} \left(u_i - g_i^{(j,k)} \right)^2$$

with $g_i^{(j,k)}$ similar to equation 8 (just without the hat symbols).

- Step 1 (initialization). Start with the no hedge situation where all β coefficients are set to zero. Application of the base procedure yields the first function estimate for step size $0 < \nu \leq 1$

$$\hat{f}^{(0)}(\cdot) = \nu \cdot \hat{g}^{(\hat{j},\hat{k})}(Y)$$

  where the replication target from equation (5) is used as pseudo-response variable $u = Y$. Set $m = 0$.

- Step 2 (update). Apply the base function to the new residuals and update the previous function estimate

$$\hat{f}^{(m+1)}(\cdot) = \hat{f}^{(m)}(\cdot) + \nu \cdot \hat{g}^{(\hat{j},\hat{k})}\left(u^{(m)}\right)$$

  with new residual vector $u^{(m)} = Y - \hat{f}^{(m)}(\cdot)$.

- Step 3 (iteration). Increase the iteration index m by one and repeat step 2 until a stopping iteration M is achieved.

4.3 COEFFICIENTS FOR US VENTURE CAPITAL FUNDS

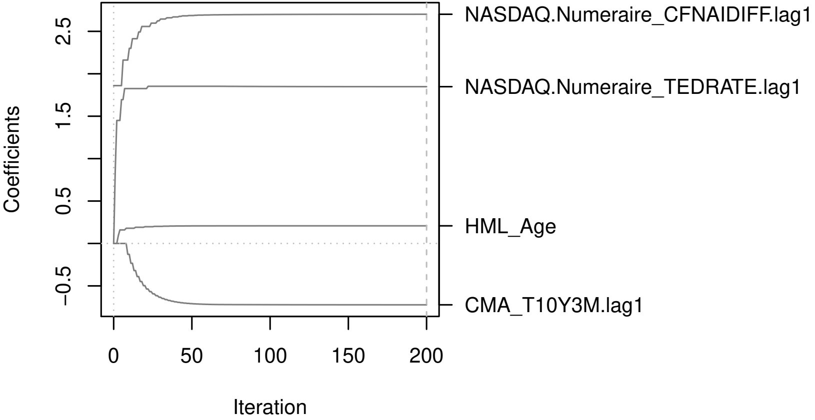

Figure: Sparse model coefficients based on stability selection for $\lambda^* = 2\%$ and $S^* = \mathrm{World}$.

4.5 HEDGE FOR US VENTURE CAPITAL FUNDS 1992

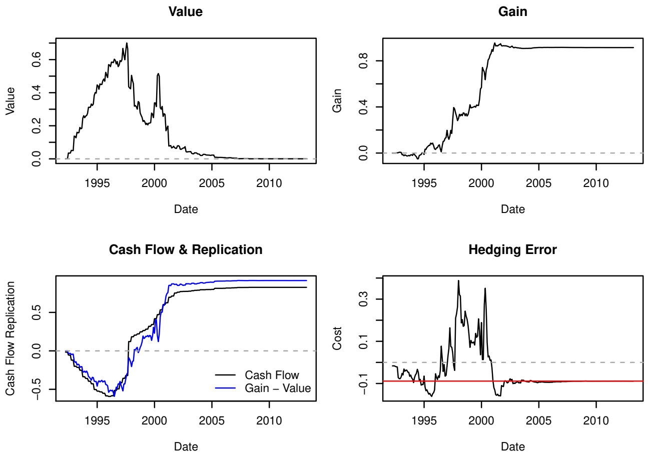

Figure: Sparse model replication strategy for US VC funds of vintage 1992. Value is V, Gain is G, Cash Flow is A, and Hedging Error is calculated as $A - (G - V) = Y - G$.

5 CONCLUSION

We presented our idea how to empirically apply dynamic hedging.

- Dynamic hedging approach theoretically more intuitive than static semiparameteric SDF estimation
- GOP numeraire portfolio is unknown. How good is your approximation? Martingale assumption very strong.
- Dynamic predictors increase likelihood of overfitting
- Even for static factors it is hard to identify significant ones (in the semiparametric papers)
- componentwise boosting and stability selection are computationally expensive

Dynamic hedging is possible but not straightforward.

- In discrete time, we don't need $\mathbb{F}$-predictability of trading strategies δ, ϑ, φ in a unique manner (before/after rebalancing)?
- Closedness of the investment opportunity set $G_T(\Theta)$ in $L^2(\mathbb{P})$: in other words, every payoff H is attainable?
- Why (sometimes) not discount H in Mean-Variance Hedging?

References

Bühlmann, H. and Platen, E. (2003).
A discrete time benchmark approach for insurance and finance.
Astin Bulletin, 33(2):153–172.

Kaplan, S. and Schoar, A. (2005).
Private equity performance: Returns, persistence, and capital flows.
Journal of Finance, 60(4):1791–1823.

Moller, T. (2001).
Risk-minimizing hedging strategies for insurance payment processes.
Finance and Stochastics, 5(4):419–446.

Pansera, J. (2012).
Discrete-time local risk minimization of payment processes and applications to equity-linked life-insurance contracts.
Insurance: Mathematics and Economics, 50:1–12.

Schweizer, M. (1995).
Variance-optimal hedging in discrete time.
Mathematics of Operations Research, 20(1):1–32.

Tausch, C. (2019).
Quadratic hedging strategies for private equity fund payment streams.
Journal of Finance and Data Science, 5(3):127–139.

Working paper and R code

will be available on my blog

Quant-Unit.com

.

Do you have comments?