

Portfolio Management and Optimal Execution via Convex Optimization

Enzo Busseti

Stanford University

April 9th, 2018

Problems

- ▶ **portfolio management**

- ▶ choose trades with optimization
- ▶ minimize risk, maximize return, control costs, ...
- ▶ *Portfolio selection*, H. Markowitz, Journal of Finance, 1952
- ▶ *A new interpretation of information rate*, J. Kelly, Bell Labs, 1956

- ▶ **optimal execution**

- ▶ execute trades in markets
- ▶ market impact
- ▶ *Optimal execution of portfolio transactions*, R. Almgren and N. Chriss, Journal of Risk, 2001

Techniques

- ▶ **convex optimization**

- ▶ 'street fighting mathematics'
- ▶ local conditions, global optimality
- ▶ *Convex Optimization*, S. Boyd and L. Vandenberghe, Cambridge University Press, 2004

- ▶ **model predictive control**

- ▶ use forecasts updated each period
- ▶ optimize trajectory of actions, use first
- ▶ *Model predictive control design: New trends and tools.*, A. Bemporad, IEEE Conference on Decision and Control, 2006

This talk

- ▶ *Multi-Period Trading via Convex Optimization*, S. Boyd, E. Busseti, S. Diamond, *et al.*, Foundations & Trends in Optimization, 2017
 - ▶ single- (and multi-) period optimization
 - ▶ coherent framework
 - ▶ risk models, costs, constraints, . . .
- ▶ *Risk-Constrained Kelly Gambling*, E. Busseti, E. Ryu, S. Boyd, Journal of Investing, 2016
 - ▶ classic problem of optimal wealth growth, with modern techniques
 - ▶ control risk of drawdown

Outline

Single-period optimization

Risk-constrained Kelly gambling

The problem

- ▶ manage a portfolio of assets over multiple periods
- ▶ take into account
 - ▶ market returns
 - ▶ transaction cost
 - ▶ holding cost
 - ▶ constraints on trades and positions
- ▶ at each period, find trades as solutions of an optimization problem
 - ▶ use up-to-date forecasts of unknown quantities
- ▶ goal is to achieve high (net) return, low risk

Definitions

- ▶ n assets, long or short, plus cash
- ▶ time periods $t = 1, \dots, T$ (e.g., trading days)
- ▶ portfolio **weights** $w_t \in \mathbf{R}^{n+1}$
 - ▶ $\mathbf{1}^T w_t = 1$
 - ▶ leverage $\|(w_t)_{1:n}\|_1$
- ▶ portfolio **value** $v_t \in \mathbf{R}$
- ▶ **holdings** $w_t v_t$
- ▶ dollar **trades** $u_t \in \mathbf{R}^{n+1}$
- ▶ **normalized trades** $z_t = u_t / v_t$
 - ▶ turnover is $\|(z_t)_{1:n}\|_1 / 2$
- ▶ asset (and cash) returns $r_t \in \mathbf{R}^{n+1}$

Multi-period trading model

- ▶ enter time period t with holdings $w_t v_t$
- ▶ choose trades z_t , execute immediately
- ▶ post-trade holdings $(w_t + z_t)v_t$
- ▶ hold for period t with market returns r_t
- ▶ next period portfolio:

$$w_{t+1}v_{t+1} = (\mathbf{1} + r_t) \circ (w_t + z_t)v_t$$

where \circ is elementwise multiplication

Transaction and holding cost

- ▶ **normalized transaction cost** $\phi_t^{\text{trade}}(z_t)$
 - ▶ (cost in dollars)/ v_t
- ▶ **normalized holding cost** $\phi_t^{\text{hold}}(z_t)$
- ▶ **self-financing condition**

$$\mathbf{1}^T z_t + \phi_t^{\text{trade}}(z_t) + \phi_t^{\text{hold}}(w_t + z_t) = 0$$

- ▶ cash balance $(z_t)_{n+1}$ determined by holdings v_t , $(w_t)_{1:n}$, and trades $(z_t)_{1:n}$

Single asset transaction cost model

- ▶ trading dollar amount x of some asset incurs cost

$$a|x| + b|x|\sigma \left(\frac{|x|}{V} \right)^{1/2} + cx$$

- ▶ a, b, c are model parameters
 - ▶ σ is one-period volatility
 - ▶ V is one-period market volume
- ▶ standard model used in practice
- ▶ variations: quadratic term, piecewise-linear, ...
- ▶ same formula for normalized trades, with $V \mapsto V/v_t$
- ▶ *Econophysics: Master curve for price-impact function*, F. Lillo, J. Farmer, and R. Mantegna, Nature, 2003

Single asset holding cost model

- ▶ holding x costs $s(x)_- = s \max\{-x, 0\}$
- ▶ $s > 0$ is shorting cost rate
- ▶ variations: quadratic term, piecewise-linear, ...
- ▶ same formula for portfolio weights

Investment

- ▶ **portfolio return** over period t :

$$R_t^p = \frac{v_{t+1} - v_t}{v_t}$$

- ▶ from the trading equation:

$$v_{t+1}/v_t = (\mathbf{1} + r_t)^T (w_t + z_t),$$

- ▶ use self-financing condition
- ▶ portfolio return in terms of weights and trades:

$$R_t^p = r_t^T (w_t + z_t) - \phi_t^{\text{trade}}(z_t) - \phi_t^{\text{hold}}(w_t + z_t)$$

Estimated portfolio return

$$\hat{R}_t^P = \hat{r}_t^T (w_t + z_t) - \hat{\phi}_t^{\text{trade}}(z_t) - \hat{\phi}_t^{\text{hold}}(w_t + z_t)$$

- ▶ quantities with $\hat{}$ are estimates or forecasts (with data available at start of period t)
- ▶ market return forecast \hat{r}_t is most important
- ▶ transaction cost estimates depend on estimates of bid-ask spread, volume, volatility
- ▶ holding cost is typically known

Single-period optimization problem I

- ▶ at start of period t , choose trades z_t using forecasts
- ▶ maximize \hat{R}_t^p , satisfy self-financing condition, trade constraints \mathcal{Z}_t , and hold constraints \mathcal{W}_t
- ▶ **risk measure** or regularizer $\psi_t(w_t + z_t)$, **risk aversion parameter** $\gamma^{\text{risk}} > 0$
- ▶ objective is estimated risk-adjusted return

$$\begin{aligned} & \text{maximize} && \hat{R}_t^p - \gamma^{\text{risk}} \psi_t(w_t + z_t) \\ & \text{subject to} && z_t \in \mathcal{Z}_t, \quad w_t + z_t \in \mathcal{W}_t, \\ & && \mathbf{1}^T z_t + \hat{\phi}_t^{\text{trade}}(z_t) + \hat{\phi}_t^{\text{hold}}(w_t + z_t) = 0 \end{aligned}$$

Single-period optimization problem II

$$\begin{aligned} & \text{maximize} && \hat{r}_t^T (w_t + z_t) \\ & && - \gamma^{\text{risk}} \psi_t(w_t + z_t) \\ & && - \hat{\phi}_t^{\text{trade}}(z_t) \\ & && - \hat{\phi}_t^{\text{hold}}(w_t + z_t) \\ & \text{subject to} && \mathbf{1}^T z_t = 0, \quad z_t \in \mathcal{Z}_t, \quad w_t + z_t \in \mathcal{W}_t \end{aligned}$$

- ▶ self-financing constraint can be approximated as $\mathbf{1}^T z_t = 0$ (slightly over-estimates updated cash balance)
- ▶ a **convex optimization problem**, if risk, trade, and hold functions/constraints are convex

Traditional quadratic risk measure

- ▶ $\psi_t(x) = x^T \Sigma_t x$
- ▶ Σ_t is an estimate of return covariance
- ▶ factor model risk $\Sigma_t = F_t \Sigma_t^f F_t^T + D_t$
 - ▶ $F_t \in \mathbf{R}^{n \times k}$ is factor exposure matrix
 - ▶ $F_t^T w_t$ are factor exposures
 - ▶ Σ_t^f is factor covariance
 - ▶ D_t is diagonal ('idiosyncratic') asset returns
 - ▶ *Large-scale portfolio optimization*, A. Perold, Management science, 1984

Robust risk measures

- ▶ worst case quadratic risk: $\psi_t(x) = \max_{i=1, \dots, M} x^T \Sigma_t^{(i)} x$
 - ▶ $\Sigma^{(i)}$ are scenario or market regime covariances

- ▶ worst case over correlation changes:

$$\psi_t(x) = \max_{\Delta} x^T (\Sigma + \Delta) x, \quad |\Delta_{ij}| \leq \kappa (\Sigma_{ii} \Sigma_{jj})^{1/2}$$

$\kappa \in [0, 1)$ is a parameter, say $\kappa = 0.05$

- ▶ can express as

$$\psi_t(x) = x^T \Sigma x + \kappa \left(\Sigma_{11}^{1/2} |x_1| + \dots + \Sigma_{nn}^{1/2} |x_n| \right)^2$$

Return forecast risk

- ▶ forecast uncertainty: any return forecast of form

$$\hat{r} + \delta, \quad |\delta| \leq \rho \in \mathbf{R}^{n+1}$$

is plausible; ρ_i is forecast return spread for asset i

- ▶ worst case return forecast is

$$\min_{|\delta| \leq \rho} (\hat{r}_t + \delta)^T (w_t + z_t) = \hat{r}_t^T (w_t + z_t) - \rho^T |w_t + z_t|$$

- ▶ same as using nominal return forecast, with a return forecast risk term $\psi_t(x) = \rho^T |x|$

Holding constraints

long only

$$w_t + z_t \geq 0$$

leverage limit

$$\|(w_t + z_t)_{1:n}\|_1 \leq L^{\max}$$

capitalization limit

$$(w_t + z_t) \leq \delta C_t / v_t$$

weight limits

$$w^{\min} \leq w_t + z_t \leq w^{\max}$$

minimum cash balance

$$(w_t + z_t)_{n+1} \geq c_{\min} / v_t$$

factor/sector neutrality

$$(F_t)_i^T (w_t + z_t) = 0$$

liquidation loss limit

$$T^{\text{liq}} \hat{\phi}_t^{\text{trade}}((w_t + z_t) / T^{\text{liq}}) \leq \delta$$

concentration limit

$$\sum_{i=1}^K (w_t + z_t)_{[i]} \leq \omega$$

Trading constraints

turnover limit $\|(z_t)_{1:n}\|_1/2 \leq \delta$

limit to trading volume $|(z_t)_{1:n}| \leq \delta(\hat{V}_T/v_t)$

transaction cost limit $\hat{\phi}^{\text{trade}}(z_t) \leq \delta$

Convexity

- ▶ objective terms and constraints above are convex, as are many others
- ▶ consequences of convexity: we can
 - ▶ (globally) solve, reliably and fast
 - ▶ add many objective terms and constraints
 - ▶ rapidly develop using domain-specific languages
- ▶ nonconvexities are easily handled, e.g.,
 - ▶ quantized positions
 - ▶ minimum trade sizes
 - ▶ target leverage (e.g., $\|(x_t + w_t)_{1:n}\|_1 = L^{\text{tar}}$)

Single-period optimization problem III

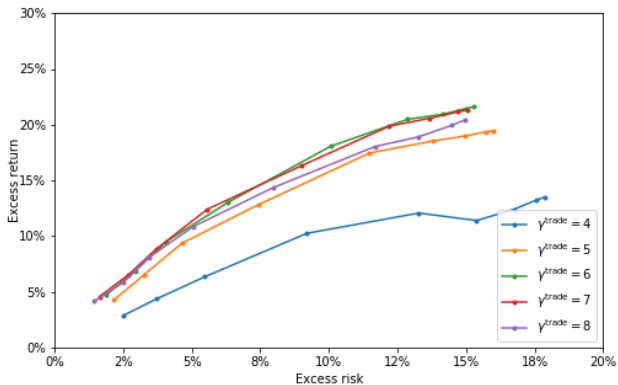
$$\begin{aligned} & \text{maximize} && \hat{r}_t^T (w_t + z_t) \\ & && -\gamma^{\text{risk}} \psi_t(w_t + z_t) \\ & && -\gamma^{\text{trade}} \hat{\phi}_t^{\text{trade}}(z_t) \\ & && -\gamma^{\text{hold}} \hat{\phi}_t^{\text{hold}}(w_t + z_t) \\ & \text{subject to} && \mathbf{1}^T z_t = 0, \quad z_t \in \mathcal{Z}_t, \quad w_t + z_t \in \mathcal{W}_t \end{aligned}$$

- ▶ $\gamma^{\text{risk}}, \gamma^{\text{trade}}, \gamma^{\text{hold}} \geq 0$ are **hyper-parameters**
- ▶ vary them to get what we want, e.g., $\hat{\phi}^{\text{trade}}$
 - ▶ absolute value term discourages small trades
 - ▶ 3/2-power term discourages large trades
 - ▶ discourages entering illiquid positions

Example

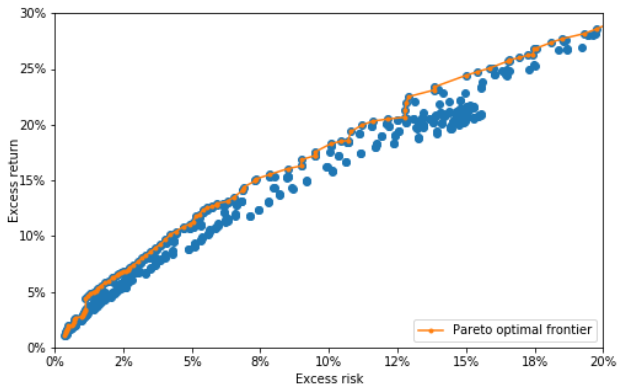
- ▶ S&P 500, daily realized returns and volumes, 2012–2016
- ▶ initial value $v_1 = \$100\text{M}$, weights $w_1 = \mathbf{1}$
- ▶ simulated market return forecasts, correct sign $\sim 54\%$ of the times
- ▶ empirical factor risk model with 15 factors
- ▶ volume, volatility forecasts are averages of last 10 values
- ▶ vary hyper-parameters γ^{risk} , γ^{trade} , γ^{hold} over ranges

Example: Risk-return trade-off



Example: Pareto optimal frontier

- ▶ grid search over 410 hyper-parameter combinations



Single-period optimization: conclusions

- ▶ organized way to parametrize trading strategies
- ▶ we defined a coherent framework
- ▶ we developed novel risk models, constraints

Outline

Single-period optimization

Risk-constrained Kelly gambling

The problem

- ▶ maximize wealth growth in repeated bets
- ▶ random i.i.d. returns
- ▶ we revisit a classic problem
 - ▶ standard convex optimization for finite outcomes
 - ▶ stochastic first order method for general returns
- ▶ standard solution can have bad drawdown
 - ▶ we develop a systematic way to trade off growth rate and drawdown risk

Model

- ▶ time periods $t = 1, 2, \dots$
- ▶ n possible **bets** (last one is cash)
- ▶ wealth $w_t \in \mathbf{R}_+$, $w_1 = 1$
- ▶ at each period, bet (same) fractions $b \in \mathbf{R}^n$ of wealth
- ▶ $\mathbf{1}^T b = 1$, $b \geq 0$
- ▶ bets have random return $r_t \in \mathbf{R}_+^n$, i.i.d., with $(r_t)_n = 1$
- ▶ wealth changes as

$$w_{t+1} = w_t r_t^T b$$

- ▶ $\log w_t = \sum_{\tau=1}^{t-1} \log r_\tau^T b$ is a random walk
- ▶ $\mathbf{E} \log(r^T b) \geq 0$ is the (wealth) **growth rate**

Kelly gambling

$$\begin{aligned} & \text{maximize} && \mathbf{E} \log(r^T b) \\ & \text{subject to} && \mathbf{1}^T b = 1, \quad b \geq 0 \end{aligned}$$

- ▶ a convex problem
 - ▶ analytical solution in trivial cases
 - ▶ standard convex optimization if outcomes are finite
 - ▶ stochastic gradient descent for general returns
- ▶ can lose large amounts of wealth before eventually increasing
 - ▶ fractional Kelly gambling
- ▶ *A new interpretation of information rate*, J. Kelly, Bell Labs, 1956

Drawdown

- ▶ $W^{\min} = \inf_{t=1,2,\dots} w_t$ is (random) function of b
- ▶ a 'drawdown' happens if W^{\min} is small
- ▶ **drawdown risk** constraint:

$$\mathbf{Prob}(W^{\min} < \alpha) < \beta, \quad \alpha, \beta \in (0, 1)$$

- ▶ problem

$$\begin{aligned} & \text{maximize} && \mathbf{E} \log(r^T b) \\ & \text{subject to} && \mathbf{1}^T b = 1, \quad b \geq 0, \\ & && \mathbf{Prob}(W^{\min} < \alpha) < \beta, \end{aligned}$$

(as far as we know) non-tractable

Lemma: threshold crossing random walk

- ▶ $(X_i, i \geq 1)$, i.i.d., with distribution measure $\mu_X(x)$
- ▶ assume $\gamma(\lambda) = \log \mathbf{E} e^{-\lambda X}$ is finite for $\lambda \in [\lambda_-, \lambda_+]$
- ▶ 'tilted' distribution $\nu_{X,\lambda}(x) = \mu_X(x)e^{-\lambda x - \gamma(\lambda)}$
- ▶ $S_n = \sum_{i=1}^n X_i$
- ▶ $\tau = \inf\{t \geq 1 | S_t < \delta\}$
- ▶ for any $n \geq 1$: $\mathbf{Prob}_{\nu_{X,\lambda}}(\tau = n) = \mathbf{E} e^{-\lambda S_\tau - \tau \gamma(\lambda)} I_{\tau=n}$
- ▶ sum over n :

$$1 \geq \mathbf{E} e^{-\lambda S_\tau - \tau \gamma(\lambda)} I_{\tau < \infty} \quad (1)$$

- ▶ similar to the *Wald's identity*, see, e.g., *Stochastic Processes: Theory for Applications*, R. Gallager, Cambridge University Press, 2013.

Drawdown bound

- ▶ $X_i = \log r_i^T b$
- ▶ $S_t = \sum_{i=1}^t \log r_i^T b = \log w_t$
- ▶ $\delta = \log \alpha$, so $\tau = \inf\{t \geq 1 | w_t < \alpha\}$
- ▶ $\lambda = \log \beta / \log \alpha > 0$
- ▶ assume $\log \mathbf{E}(r^T b)^{-\lambda} = \gamma(\lambda) \leq 0$
- ▶ $S_\tau < \log \alpha$ and $\tau\gamma(\lambda) \leq 0$, so from (1):

$$1 \geq \mathbf{E} e^{-\lambda S_\tau - \tau\gamma(\lambda)} I_{\tau < \infty} > \mathbf{E} e^{-\lambda \log \alpha} I_{\tau < \infty}$$

- ▶ which implies:

$$\beta > \mathbf{Prob}(W^{\min} < \alpha)$$

Risk-constrained Kelly gambling

$$\begin{aligned} & \text{maximize} && \mathbf{E} \log(r^T b) \\ & \text{subject to} && \mathbf{1}^T b = 1, \quad b \geq 0 \\ & && \log \mathbf{E}(r^T b)^{-\lambda} \leq 0, \end{aligned}$$

- ▶ disciplined convex problem
- ▶ **convex restriction** of risk constrained problem
- ▶ solution satisfies $\mathbf{Prob}(W^{\min} < \alpha) < \beta$, for any $\alpha, \beta \in (0, 1)$ such that $\lambda = \log \beta / \log \alpha$

Example

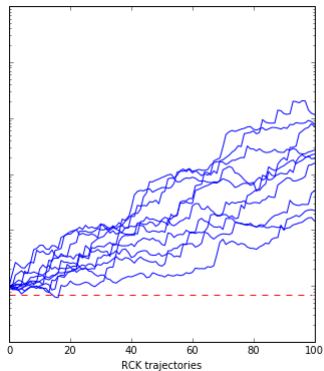
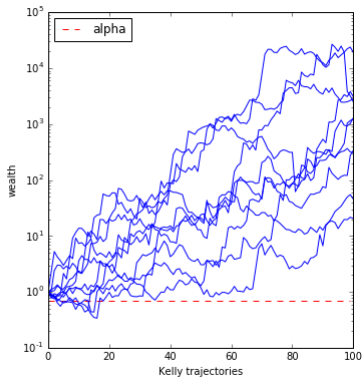
- ▶ $n = 20$, $K = 100$ outcomes
- ▶ probabilities $\pi_i \geq 0$, $i = 1, \dots, K$
- ▶ returns $r_{ij} \in \mathbf{R}_{++}$ for $i = 1, \dots, K$ and $j = 1, \dots, n - 1$, uniform in $[0.7, 1.3]$, plus some 'outliers'
- ▶ $\alpha = 0.7$, $\beta = 0.1$, $\lambda = 6.456$
- ▶ solve problem
- ▶ simulate trajectories of w_t for $t = 1, \dots, T = 100$

Bet	$\mathbf{E} \log(r^T b)$	$e^{\lambda \log \alpha}$	$\mathbf{Prob}(W^{\min} < \alpha)$
Kelly	0.062	-	0.398
RCK, $\lambda = 6.456$	0.043	0.100	0.071
RCK, $\lambda = 5.500$	0.047	0.141	0.099

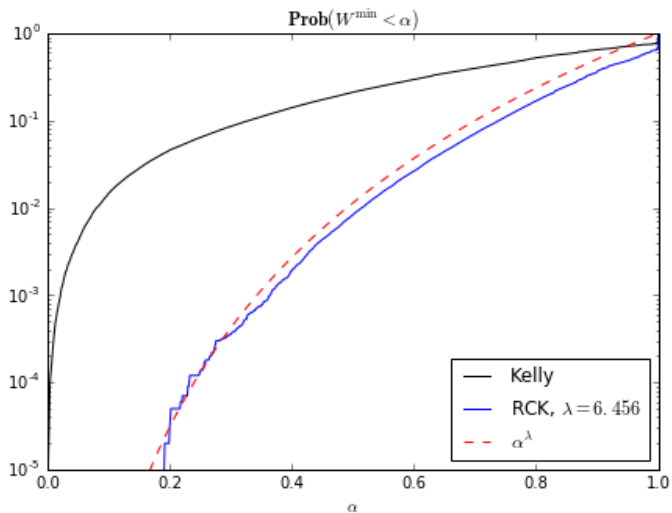
Example: CVXPY code

```
b = Variable(n)
lambd = Parameter(sign = 'positive')
growth = pi.T*log(r.T*b)
risk = log_sum_exp(log(pi) - lambd * log(r.T*b)) <= 0
constr = [sum_entries(b) == 1, b >= 0, risk]
Problem(Maximize(growth), constr).solve()
```

Example: wealth trajectories



Example: CDF of W^{\min}



Risk-constrained Kelly gambling: conclusions

- ▶ Kelly gambling maximizes growth, but can have large drawdown
- ▶ we developed
 - ▶ a systematic way to trade off growth rate and drawdown risk
 - ▶ convex optimization algorithms to solve the problem with general returns

Thank you!

Quadratic approximation I

- ▶ assume $r^T b \approx 1$, and let $\rho = r - \mathbf{1}$:

$$\log(r^T b) = \rho^T b - \frac{1}{2}(\rho^T b)^2 + O((\rho^T b)^3)$$

$$(r^T b)^{-\lambda} = 1 - \lambda \rho^T b + \frac{\lambda(\lambda + 1)}{2}(\rho^T b)^2 + O((\rho^T b)^3)$$

- ▶ we get:

$$\begin{aligned} & \text{maximize} && \mathbf{E} \rho^T b - \frac{1}{2} \mathbf{E}(\rho^T b)^2 \\ & \text{subject to} && \mathbf{1}^T b = 1, \quad b \geq 0, \\ & && -\lambda \mathbf{E} \rho^T b + \frac{\lambda(\lambda+1)}{2} \mathbf{E}(\rho^T b)^2 \leq 0 \end{aligned}$$

- ▶ quadratically constrained (convex) quadratic program
- ▶ equivalent to Markowitz problem with $\mu = \mathbf{E} \rho$,
 $\Sigma = \mathbf{E} \rho \rho^T - \mu \mu^T$

Quadratic approximation II

- ▶ dualize the risk constraint, for some $\nu \geq 0$

$$\begin{aligned} & \text{maximize} && \mu^T b - \frac{1}{2} b^T S b + \nu (\mu^T b - \frac{\lambda+1}{2} b^T S b) \\ & \text{subject to} && \mathbf{1}^T b = 1, \quad b \geq 0 \end{aligned}$$

- ▶ rearrange and scale constants

$$\begin{aligned} & \text{maximize} && \mu^T b - \frac{\eta}{2} (\mu^T b)^2 - \frac{\eta}{2} b^T \Sigma b \\ & \text{subject to} && \mathbf{1}^T b = 1, \quad b \geq 0, \end{aligned}$$

- ▶ equivalent to

$$\begin{aligned} & \text{maximize} && (1 - \eta \mu^T b^*) \mu^T b - \frac{\eta}{2} b^T \Sigma b \\ & \text{subject to} && \mathbf{1}^T b = 1, \quad b \geq 0, \end{aligned}$$

(can show that $\mu^T b^* \geq 1/\eta$)

Example: Trade-off of drawdown risk and growth rate

