Recitation 5

Friday, March 19, 2021 10:35 AM

QUESTIONS TO ANSWER TODAY

Why would we choose a certain uncertainty set? At a high level, because:

- It represents underlying data. (Portfolio optimization)
- It is less conservative compared to another set, with similar guarantees. (Entropy uncertainty set.)

General purpose method for deriving robust counterpart (HW2 Q 1.1). High level:

- Divide the problem into more manageable subproblems.
- Use properties of the conjugate.
- Instead of trying to take the conjugate of the larger problems, take conjugate of subproblems by relating to functions you know.

Talking about the problem set code, and coding tips.

These notes will be posted.

Restricted Brownian Motion in Portfolio Optimization

Friday, March 19, 2021 10:26 AM

(Lecture 7 on Port. Opt., slide 4)

The deterministic problem

$$
\begin{aligned}
\text{Max} \quad & \sum_{m=0}^{M} x_N^m \\
\text{s.t.} \quad & x_t^m = (1+\tilde{r}_{t-1}^m)(x_{t-1}^m - u_{t-1}^m + v_{t-1}^m) \\
& x_t^0 = (1+r_{t-1}^0)\left(x_{t-1}^0 + \sum_{m=1}^{M}(1-c_{sell})u_{t-1}^m - \sum_{m=1}^{M}(1+c_{buy})v_{t-1}^m\right) \\
& x_t^m \geq 0 \\
& u_t^m \geq 0,\ v_t^m \geq 0
\end{aligned}
$$

Reformulated with auxiliary variables

$$
\begin{aligned}
R_0^m &= 1 \\
\underline{\tilde{R}_t^m} &= (1+\tilde{r}_0^m)(1+\tilde{r}_1^m)\dots(1+\tilde{r}_{t-1}^m) \\
\xi_t^m &= \frac{x_t^m}{\tilde{R}_t^m}, \quad \eta_t^m = \frac{u_t^m}{\tilde{R}_t^m}, \quad \zeta_t^m = \frac{v_t^m}{\tilde{R}_t^m}
\end{aligned}
$$

Why? $\Sigma_1 \rightarrow$ variance of 1-period returns

$\Sigma_t \rightarrow$ " " t-period " .

We are able to better use data.

Uncertainty set choice?

- Uncertainty sets:

norm-bounded on 1st timestep,

$$
U_1 = \left\{ \begin{array}{l}
\|\Sigma_1^{-\frac{1}{2}}\left(\tilde{\boldsymbol{R}}_1 - \overline{\boldsymbol{R}}_1\right)\| \leq \Delta \\
\underline{\delta}_2^m \tilde{R}_1^m \leq \tilde{R}_2^m \leq \overline{\delta}_2^m \tilde{R}_1^m, \ m = 1, \dots, M \\
\vdots \\
\underline{\delta}_N^m \tilde{R}_{N-1}^m \leq \tilde{R}_N^m \leq \overline{\delta}_N^m \tilde{R}_{N-1}^m, \ m = 1, \dots, M
\end{array} \right\}
$$

then bounded by polyhedron

where $\underline{\delta}_t^m$ and $\overline{\delta}_t^m$ are of the form $(1+\underline{r}_t^m)$ and $(1+\overline{r}_t^m)$, respectively.

- If data on the covariance matrices of future cumulative returns are available:

where $\underline{\delta}_t^m$ and $\overline{\delta}_t^m$ are of the form $(1 + \underline{r}_t^m)$ and $(1 + \overline{r}_t^m)$, respectively.

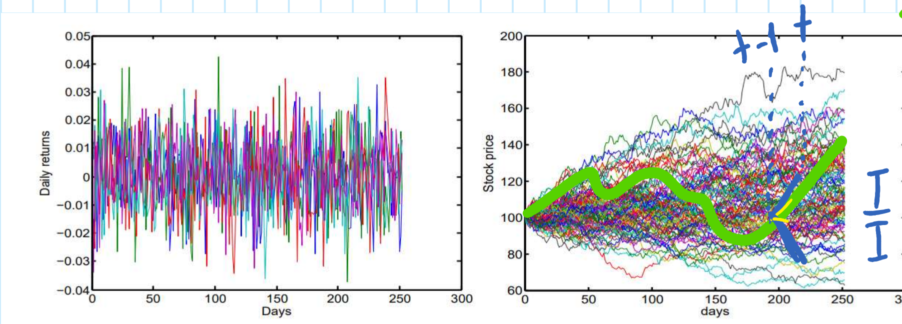
- If data on the covariance matrices of future cumulative returns are available:

$$U_2 = \left\{ \ ||\Sigma_t^{-\frac{1}{2}} \left(\tilde{\boldsymbol{R}}_t - \overline{\boldsymbol{R}}_t \right) || \leq \Delta_t, \ t = 1, \dots, N \ \right\},$$

bounded by N norm sets

where Δ_t are budgets of uncertainty.

Why? Because it represents some underlying structure in the data!



some potential stock price trajectory

$1 + \overline{r}_{t-1}^m = \overline{\delta}_{t-1}^m$

$1 + \underline{r}_{t-1}^m = \underline{\delta}_{t-1}^m$

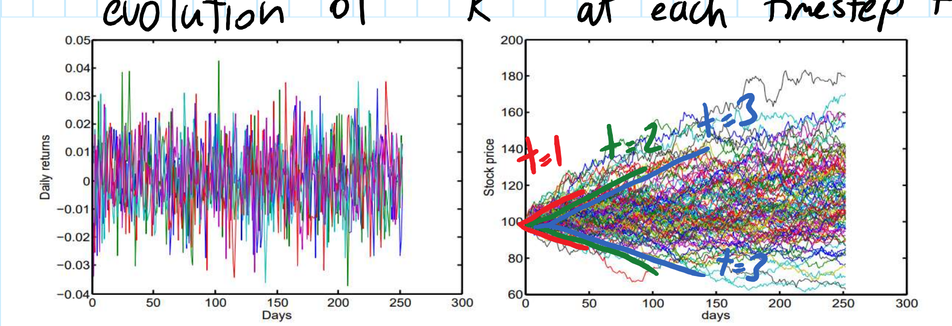
some bounded polyhedral evolution

First case:

$$\underline{\delta}_{t-1}^m \tilde{R}_{t-1}^m \leq \tilde{R}_t^m \leq \overline{\delta}_{t-1}^m \tilde{R}_{t-1}^m$$

$\underline{\delta}_{t-1}^m \to 1 + \underline{r}_{t-1}^m$ $\qquad$ $\overline{\delta}_{t-1}^m \to 1 + \overline{r}_{t-1}^m$

The polyhedron describes some uncertain cone of evolution of $\tilde{R}^m$ at each timestep t.



Second case:

$$\left\| \Sigma_t \left(\tilde{R}_t^m - \overline{R}_t^m \right) \right\| \leq \Delta_t$$

The set describes another cone of evolution, but based on the $t=0$ time instead of $t-1$.

Entropy Uncertainty Set

Thursday, March 18, 2021 02:52 PM

Why is the entropy uncertainty set useful?

Example 2.9. Entropy uncertainty set. *In [222] a so-called entropy uncertainty set is derived that is tighter than the ball-box uncertainty set, but still has the same probability bound. This set is defined as*

$$\mathcal{Z} = \left\{ \mathrm{z} \;\middle|\; \|\mathrm{z}\|_\infty \le 1,\ \sum_{\ell=1}^{L} \{(1+z_\ell)\ln(1+z_{ell}) + (1-z_{ell})\ln(1-z_\ell)\} \le \beta \right\}.$$

Takeaway: Entropy set is useful since it provides the same probabilistic guarantees as the ball-box set, but is smaller, and thus less conservative!

In the problem set, you showed the below...

Hence, using Lemma 2.1, (RC) can be written as

$$\bar{\mathrm{a}}^T \mathrm{x} + \|\mathrm{w}\|_\infty + \sum_{\ell=1}^{L} \left\{ t_\ell - s_\ell + u\left(e^{s_\ell/u-1} + e^{-t_\ell/u-1}\right)\right\} \le b,$$
$$\mathrm{w} + \mathrm{s} + \mathrm{t} = \mathrm{P}^T \mathrm{x},$$
$$\mathrm{u} \ge \mathbf{0}.$$

Firstly, what does this result actually mean? Is it useful?

Just because you can take RC, doesn't mean it's useful, since I don't know how to solve the RC.

Motivation for Ball-Box uncertainty set

Suppose: z_i are stochastic and independent, with known support, say $[-1, 1]$ zero mean.

Consider: $Z_\Omega = \{z \;:\; \|z\|_2 \le \Omega\}$ (ball)

$\|z\|_\infty \le 1$

$$(RC) \quad (a + Pz)^T x \le b, \ \forall z \in Z_\Omega.$$

Remember ball-box set guarantees.

If x satisfies (3) then:

$$\text{Prob}\left((a + Pz)^T x \le b\right) \ge 1 - \exp(-\Omega^2/2) =: 1 - \epsilon.$$

So: $\quad \Omega = \sqrt{2\ln(1/\epsilon)}$.

Example 1: $\Omega = 7.44 \Longrightarrow \epsilon = 10^{-12}$.

Entropy set has the same guarantees, when

So: $\Omega = \sqrt{2\ln(1/\epsilon)}$.

Entropy set has the same guarantees, when the safety factor is $\beta = \Omega^2$.

Example 1: $\Omega = 7.44 \Longrightarrow \epsilon = 10^{-12}$.
Example 2: $\Omega = 3.255 \Longrightarrow \epsilon = 0.005$.
Example 3: $\Omega = 2.444 \Longrightarrow \epsilon = 0.05$.

Same result for $Z_\Omega = \{z \,:\, \|z\|_2 \le \Omega,\ \|z\|_\infty \le 1\}$ (ball-box).

I'm not going to show why it has the same probabilistic guarantees, but I will show why it is a smaller set!

Obviously, both are still bounded by $\underline{\|z\|_\infty \le 1}$

We need to show

$$\sum_{l=1}^{L} \left(\log(..) + \dots\right) \le \beta \quad \text{is a smaller set}$$

$$\text{than} \quad \|z\|_2 \le \sqrt{\beta}.$$

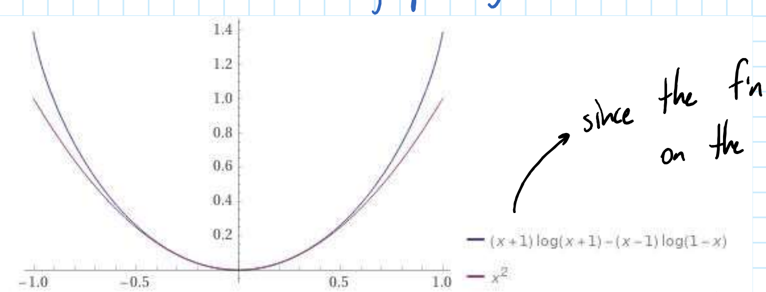
This is true iff

$$\sum_{l=1}^{L} (1+z_l)\log(1+z_l) + (1-z_l)\log(1-z_l) \ge \sum_{l=1}^{L} z_l^2$$

$$(1+z_l)\log(1+z_l) + (1-z_l)\log(1-z_l) \ge z_l^2, \quad -1 \le z_l \le 1$$

How to show? Use this fact: $x-1 \ge \ln(x) \ge 1 - \frac{1}{x}$, $0 < x < 1$

OR even better, graphically.



since the f'n is larger on the bounded domain $z_l \in [-1, 1]$, the entropy set is a subset of the ball-box set!

Where does the set come from?

[1] Nemirovski, A., "Lectures on robust convex optimization," *Class notes, Georgia*

Inst. of Technology, 2009.

Example 2.1 Let $\mathcal{P}$ be the set of all *product-type* probability distributions on $\mathbb{R}^L$ (that is, $\zeta_1, ..., \zeta_L$ are independent whenever $\zeta \sim P \in \mathcal{P}$) with zero mean marginal distributions P_ℓ supported on $[-1, 1]$. Choosing as the generator the function $\gamma(s) = \exp\{s\}$, let us compute an appropriate Ψ^+. This should be a convex function such that

$$\forall w \in \mathbb{R}^{L+1} : \Psi^+(w) \geq \mathbf{E}_{\zeta\sim P}\left\{\exp\{w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell\}\right\} \ \forall P \in \mathcal{P}.$$

Given $P \in \mathcal{P}$, denoting by P_ℓ the marginal distributions of P and taking into account that P is product-type, we get

$$\mathbf{E}_{\zeta\sim P}\left\{\exp\{w_0 + \sum_{\ell=1}^{L} \zeta_\ell w_\ell\}\right\} = \exp\{w_0\} \prod_{\ell=1}^{L} \mathbf{E}_{\zeta_\ell\sim P_\ell}\{\exp\{w_\ell \zeta_\ell\}\}.$$

Now let us use the following simple and well known fact:

Lemma 2.1 *Let Q be a zero mean probability distribution on the axis supported on $[-1, 1]$. Then for every real t one has*

$$\mathbf{E}_{s\sim Q}\{\exp\{ts\}\} \leq \cosh(t). \qquad (2.3.11)$$

Follow the proof if you are interested.

Homework 2 Part 1.1

Tuesday, March 16, 2021 10:06 AM

$$h(z) = \sum_{i=1}^{r} \frac{\alpha_i}{p_i} |(d^i)^T z - \beta_i|^{p_i}, \quad \alpha_i > 0,\ p_i > 1$$

Derive the robust counterpart of

$$(\bar{a} + Pz)^T x \le b \quad \forall\ z \in Z = \{z : h(z) \le \Gamma\}$$

Use two key concepts to find the RC of any set!

→ The conjugate of functions we know.

→ The properties of conjugates.

Reminder: Need to find:

$$\bar{a}^T x + \underbrace{\delta^*(P^T x \mid Z)}_{} \le b \longrightarrow \text{conjugate of support f'n of } P^T x \text{ on set } Z.$$

Conjugate f'n of the norm

$$\left(\frac{|t|^p}{t}\right)^*_{p>1} = \frac{|s|^q}{s} \quad \text{where} \quad \frac{1}{p} + \frac{1}{q} = 1$$

(dual norm)

Properties of conjugates (p. 39) (for $\alpha > 0$)

Let's say $h(z) = g(D^T z) = \sum_{i=1}^{r} g_i(d_i^T z)$

Page 48

$$g(D^T z) \le 0 \longrightarrow \begin{cases} \bar{a}^T x + u g^*\left(\frac{v}{u}\right) \le b \\ D^T v = P^T x \\ u \ge 0 \end{cases}$$

Let's represent $\quad g(D^T z) \le \Gamma$

$g_i(d_i^T z)$ with $g_i(x_i)$ where $x_i = d_i^T z$

$$g_i = \frac{\alpha_i}{p_i} |x_i - \beta_i|^{p_i} - \frac{\Gamma}{r} \longrightarrow \text{Want to relate this to } \frac{|t|^p}{t} \text{ norm}$$

$g_i = \frac{\alpha_i}{p_i} |x_i - \beta_i| - \frac{\ }{r} \longrightarrow$ want to relate ... to f

scaling

$(\alpha f)^*(y) = \alpha f^*\left(\frac{y}{\alpha}\right)$

sideshift

$\tilde{f}(x) = f(x-a)$

$\tilde{f}^*(y) = f^*(y) + a^T y$

downshift

General form of conjugate of $\tilde{g}$ (i.e. transformed g, for which we already know the conjugate)

$$\tilde{g} = \alpha + \beta x + \gamma \cdot g(\lambda x + \delta)$$

$$\tilde{g}^*(y) = -\alpha - \delta \frac{y - \beta}{\lambda} + \gamma \cdot g^*\left(\frac{y-\beta}{\gamma \lambda}\right) \quad (\gamma > 0)$$

Plugging in... $\alpha \to -\frac{\Gamma}{r}$, $\delta \to -\beta_i$, $\lambda \to 1$, $\beta \to 0$, $\gamma \to \alpha_i$, $\lambda \to 1$, $\beta \to 0$

$$\tilde{g}^*_i(y) = \frac{\Gamma}{r} + \beta_i^T y + \alpha_i \, g^*\left(\frac{y}{\alpha_i}\right)$$

Apply $\left(\frac{|t|^p}{t}\right)^*(y) = \frac{|y|^q}{q}$

$$\tilde{g}^*_i(z) = \frac{\Gamma}{r} + \beta_i^T z + \frac{\alpha_i}{q_i}\left|\frac{z_i}{\alpha_i}\right|^q$$

Remember:

$$g(D^T z) \le 0 \longrightarrow \begin{cases} \bar{a}^T x + u g^*\left(\frac{v}{u}\right) \le b \\ Dv = P^T x \\ u \ge 0 \end{cases}$$

Thus the safety factor is

$$u \sum_{i=1}^{r} \left(\frac{\Gamma}{r} + \beta_i^T\left(\frac{z}{u}\right) + \frac{\alpha_i}{q_i}\left|\frac{z_i}{\alpha_i u}\right|^q\right) = \Gamma u + \beta^T z + \sum_{i=1}^{r} \frac{\alpha_i}{q_i}\left|\frac{z_i}{\alpha_i u}\right|^q$$

$$u \sum_{i=1}^{r} \left(\frac{1'}{r} + \beta_i' \left(\frac{z}{u}\right) + \frac{\alpha_i}{q_i} \left| \frac{z_i}{\alpha_i u} \right| \right) = 1'u + \beta' z + \sum_{i=1} \frac{\quad}{q_i} \left| \frac{\quad}{\alpha_i u} \right|$$

And the robust counterpart

$$\bar{a}^T x + u \sum_{i=1}^{r} \left(\frac{\alpha_i}{q_i}\right) \left| \frac{z_i}{\alpha_i u} \right|^{q_i} + \beta^T z + P_u \leq b$$

$$\text{s.t.} \quad D^T z = P^T x, \quad u \geq 0, \quad \frac{1}{p_i} + \frac{1}{q_i} = 1, \ \forall i$$