Dynamic programming for mean–variance portfolio selection

Martin Schweizer

Department of Mathematics, ETH Zürich,
and Swiss Finance Institute

Conference in memory of Tomas Björk
Swedish House of Finance, Stockholm, October 10-11, 2022

Stockholm, October 10, 2022

based on joint work with **Zhouyi Tan**

The basic problem

- Consider **financial market** with discounted prices given by bank account 1 and stocks S, where S is $\mathbb{R}^d$-valued semimartingale.
- Call $G(\vartheta) = \int \vartheta \, \mathrm{d}S$ the **gains from trade** for self-financing strategy with initial capital 0 and using ϑ to trade in S.
- **MVPS problem** is to

$$\text{maximise } E[G_T(\vartheta)] - \xi \text{Var}[G_T(\vartheta)] \text{ over all } \vartheta \in \Theta$$

  for some risk-aversion parameter $\xi > 0$.
- We want to find an optimal ϑ^* via **dynamic programming**.

Abstract setup

- $\mathcal{X} = \{X^\vartheta : \vartheta \in \Theta\}$ family of processes on $\{0, 1, \dots, T\}$.
- Θ is family of predictable processes (controls) ϑ.
- **Controlled:** $X^\vartheta|_{[\![0,k]\!]}$ only depends on $\vartheta|_{]\!]0,k]\!]}$ for each k.
- **Goal** is to maximise

$$j(\vartheta) := E\Bigg[\sum_{k=1}^{T} f\big(k, X^\vartheta|_{[\![0,k]\!]}, \vartheta|_{]\!]0,k]\!]}, \mathcal{L}(X^\vartheta|_{[\![0,k]\!]}, \vartheta|_{]\!]0,k]\!]})\big) + g\big(X^\vartheta|_{[\![0,T]\!]}, \mathcal{L}(X^\vartheta|_{[\![0,T]\!]})\big)\Bigg]$$

over all $\vartheta \in \Theta$, with suitable functions f **(running reward)** and g **(final reward)**.

Examples

- **Example 1.** If f and g do not depend on the law $\mathcal{L}$, this is a **classical stochastic control problem**.
- **Example 2.** If $f \equiv 0$ and g is affine–quadratic in X_T^ϑ and in $E[X_T^\vartheta] = \int z\mathcal{L}(X_T^\vartheta)(\mathrm{d}z)$, we can get as a special case the **MVPS problem**.
- **Example 3.** If f has the form

$$f\big(k, X^\vartheta|_{[\![0,k]\!]}, \vartheta_k, \mathcal{L}(X^\vartheta|_{[\![0,k]\!]}, \vartheta_k)\big) = -\ell(\vartheta_k \Delta X_k^\vartheta - E[\vartheta_k \Delta X_k^\vartheta])$$

for a loss function ℓ, we want to avoid strategies whose instantaneous gains fluctuate too much.

- Define $\Theta(k,\vartheta) := \{\vartheta' \in \Theta : \vartheta' = \vartheta \text{ on }]\!]0,k]\!]\}$,

$$j(k,\vartheta') := E\Bigg[\sum_{i=k+1}^{T} f\big(i, X^{\vartheta'}|_{[\![0,i]\!]}, \vartheta'|_{]\!]0,i]\!]}, \mathcal{L}(X^{\vartheta'}|_{[\![0,i]\!]}, \vartheta'|_{]\!]0,i]\!]})\big)$$

$$+ g\big(X^{\vartheta'}|_{[\![0,T]\!]}, \mathcal{L}(X^{\vartheta'}|_{[\![0,T]\!]})\big)\Bigg],$$

$$v(k,\vartheta) := \sup\{j(k,\vartheta') : \vartheta' \in \Theta(k,\vartheta)\}.$$

- Then

$$v(T,\vartheta) = j(T,\vartheta) = E\Big[g\big(X^{\vartheta}|_{[\![0,T]\!]}, \mathcal{L}(X^{\vartheta}|_{[\![0,T]\!]})\big)\Big],$$

$$v(0,\vartheta) = \sup\{j(\vartheta') : \vartheta' \in \Theta\} =: v(0).$$

- **Theorem:** Assume that $\Theta(\ell, \vartheta) \subseteq \Theta(k, \vartheta)$ for $\ell \geq k$. Then:

1) For $k = 0, 1, \dots, T-1$ and any $\vartheta \in \Theta$,

$$
\begin{aligned}
& v(k-1, \vartheta) \\
& = \sup_{\vartheta' \in \Theta(k-1, \vartheta)} \Big\{ v(k, \vartheta') \\
& \qquad + E\Big[f\Big(k, X^{\vartheta'}|_{[\![0,k]\!]}, \vartheta'|_{]\!]0,k]\!]}, \mathcal{L}\big(X^{\vartheta'}|_{[\![0,k]\!]}, \vartheta'|_{]\!]0,k]\!]}\big)\Big)\Big] \Big\}.
\end{aligned}
$$

2) In consequence, $v(k-1, \vartheta)$ only depends on $\vartheta|_{]\!]0,k-1]\!]}$, and the maximisation above only goes over those random variables δ_k which make up the k-th time coordinate of ϑ'.

Comments

- The **dynamic programming principle** here is for a **deterministic function**, but we still **optimise over stochastic quantities**.
- For finite discrete time, DPP leads to **backward recursion for value function**, with each recursion step a **single-step problem** where we **optimise over only one variable**.
- Because our **criterion** depends both on the state X^ϑ and on its law $\mathcal{L}(X^\vartheta)$, it **includes both**
  - **expectations of nonlinear functions** of the state (as in classical stochastic control), and
  - **nonlinear functions of expectations** of the state (as in e.g. MVPS).

A linear–quadratic example

- Suppose running reward $f \equiv 0$ and final reward is

$$g(x,\mu) = a_T x + b_T x^2 + c_T \left(\int z\,\mu(\mathrm{d}z) \right)^2 + d_T,$$

with deterministic coefficients a_T, b_T, c_T, d_T.

- So we want to maximise

$$a_T E[X_T^\vartheta] + b_T E[(X_T^\vartheta)^2] + c_T (E[X_T^\vartheta])^2 + d_T.$$

- If $X^\vartheta = \int \vartheta\,\mathrm{d}S = G(\vartheta)$ for some fixed process S, we obtain the **MVPS problem** for

$$a_T = 1, \quad b_T = -\xi, \quad c_T = \xi, \quad d_T = 0.$$

LQ control and DPP

- In the setting with $f \equiv 0$, the DPP takes the form

$$v(k-1,\vartheta) = \sup\left\{v\big(k,\vartheta(k,\delta_k)\big) : \delta_k \in \Theta^{[k]}(\vartheta)\right\}, \qquad (1)$$

where $\vartheta(k,\delta_k) = (\vartheta_1,\dots,\vartheta_{k-1},\delta_k,\vartheta_{k+1},\dots,\vartheta_T)$ and

$$\Theta^{[k]}(\vartheta) := \{\vartheta'_k : \vartheta' \in \Theta(k-1,\vartheta)\}.$$

- We also start the recursion with

$$\begin{aligned} v(T,\vartheta) &= E\left[g\big(X_T^\vartheta, \mathcal{L}(X_T^\vartheta)\big)\right] \\ &= a_T E[X_T^\vartheta] + b_T E[(X_T^\vartheta)^2] + c_T (E[X_T^\vartheta])^2 + d_T. \end{aligned}$$

- Now we study this for the case $X^\vartheta = \int \vartheta \,\mathrm{d}S = G(\vartheta)$ with a fixed process S. So dynamics of X^ϑ is **linear**.

Solving the LQ problem via DPP

- To find optimal strategy ϑ^*, we **solve each subproblem** (1) at time k to find an optimal δ_k^* there. Then $\vartheta^* = (\delta_1^*, \ldots, \delta_T^*)$.
- How do we solve (1)?
  - For $k = T$, criterion $v(T, \vartheta)$ depends on X_T^ϑ and $E[X_T^\vartheta]$ in **affine–quadratic** way.
  - We need to optimise over the last component δ_T of ϑ.
  - **Linear state dynamics** gives $X_T^\vartheta = X_{T-1}^\vartheta + \delta_T \Delta S_T$.
  - So criterion for δ_T is **affine–quadratic**, with coefficients depending on X_{T-1}^ϑ and $E[X_{T-1}^\vartheta]$ in affine–quadratic way.
  - **FOC** for optimal δ_T^* will be **linear**, again with coefficients depending on X_{T-1}^ϑ and $E[X_{T-1}^\vartheta]$ in affine–quadratic way.
  - So we **expect** that **optimal** δ_T^* will be **linear**, depending on X_{T-1}^ϑ and $E[X_{T-1}^\vartheta]$ in **affine–quadratic** way.

Solving the LQ problem via DPP (cont'd)

- Plugging δ_T^* back into $v(T-1,\vartheta) = v(T, \vartheta(T, \delta_T^*))$, we expect again affine–quadratic function of X_{T-1}^ϑ and $E[X_{T-1}^\vartheta]$.
- **This should then iterate ...**

- So we **guess** that we should obtain

$$v(k,\vartheta) = a_k E[X_k^\vartheta] + b_k E[(X_k^\vartheta)^2] + c_k (E[X_k^\vartheta])^2 + d_k,$$

with deterministic coefficients a_k, b_k, c_k, d_k.

- However, the **correct form** is actually

$$v(k,\vartheta) = a_k E[Z_k X_k^\vartheta] + b_k E[Z_k (X_k^\vartheta)^2] + c_k (E[Z_k X_k^\vartheta])^2 + d_k,$$

with deterministic coefficients a_k, b_k, c_k, d_k and with $Z_k \geq 0$, **$\mathcal{F}_k$-measurable and bounded**.

Solving the LQ problem via DPP (cont'd)

- Starting with the **ansatz**

$$v(k,\vartheta) = a_k E[Z_k X_k^\vartheta] + b_k E[Z_k (X_k^\vartheta)^2] + c_k (E[Z_k X_k^\vartheta])^2 + d_k, \tag{2}$$

  with a_k, b_k, c_k, d_k deterministic and $Z_k \geq 0$, $\mathcal{F}_k$-measurable and bounded, **one can** now
  - **iteratively** work out the **optimal** δ_k^* via the FOC,
  - express δ_k^* in terms of a_k, b_k, c_k, d_k and Z_k as well as X_{k-1}^ϑ,
  - **check recursively** that (2) holds, and
  - also obtain a **recursion for** the **coefficients** a, b, c, d and Z.

Technical issues

- Where do we encounter **difficulties** with the above procedure**?**
- One needs to check in the recursion for the coefficients that the **properties** of Z_k are **inherited** by Z_{k-1}.
- One needs to check that everything that comes up in formulas is **well defined**.
- One needs to check that the candidate δ_k^* obtained by solving the FOC is **sufficiently integrable**, so that one can later deduce that $\vartheta^* := (\delta_1^*, \dots, \delta_T^*)$ is in the correct space Θ.

- Which **assumptions** do we need**?**

  1) S is **square-integrable** and satisfies **structure condition SC**, meaning that $A \ll \langle M \rangle$, where

  $$A := \sum E[\Delta S_k | \mathcal{F}_{k-1}], \quad \langle M \rangle := \sum \mathrm{Var}[\Delta S_k | \mathcal{F}_{k-1}].$$

  2) The **closure** in $L^2(P)$ of the space

  $$G_T(\Theta) := \left\{ G_T(\vartheta) := \sum_{j=1}^{T} \vartheta_j \Delta S_j : \vartheta \in \Theta \right\}$$

  does not contain the constant payoff 1.

  3) Θ is such that $\Theta(\ell, \vartheta) \subseteq \Theta(k, \vartheta)$ for $\ell \geq k$.

- Note that both 1) and 2) are **absence-of-arbitrage** type conditions.

Example spaces

- **Possible choices** for the space Θ of **integrands/controls** ϑ:
  - We always assume that ϑ is predictable.
  - Notation: $G_k(\vartheta) := \sum_{j=1}^{k} \vartheta_j \Delta S_j$ for $k = 0, 1, \dots, T$.
  - Can consider

$$\Theta_{\mathrm{S}} := \{\text{all } \vartheta \text{ with } G_k(\vartheta) \in L^2(P) \text{ for all } k\} \qquad [\longrightarrow \text{Schweizer}]$$

$$\Theta_{\mathrm{CK}} := \{\text{all } \vartheta \text{ for which there is a sequence } (\vartheta^n)_{n\in\mathbb{N}} \text{ of elementary strategies with } G_k(\vartheta^n) \to G_k(\vartheta) \text{ in } L^0 \text{ for all } k \text{ and } G_k(\vartheta^n) \to G_k(\vartheta) \text{ in } L^2(P)\} \qquad [\longrightarrow \text{Černý/Kallsen}]$$

$$\Theta_{\mathrm{MN}} := \{\text{all } \vartheta \text{ with } G_T(\vartheta) \in L^2(P)\} \qquad [\longrightarrow \text{Melnikov/Nechaev}]$$

- **Theorem:** Under **assumptions 1)–3)**, the **optimal strategy** for the MVPS problem is **recursively** given by

$$\vartheta_k^* = -\frac{E[Z_k \Delta S_k | \mathcal{F}_{k-1}]}{E[Z_k (\Delta S_k)^2 | \mathcal{F}_{k-1}]} \left(G_{k-1}(\vartheta^*) - \frac{1}{2\xi E[Z_0]} \right).$$

- For the case $\Theta = \Theta_{\mathrm{S}}$, this needs **extra assumptions** to guarantee that ϑ^* is in Θ_{S}:
  - A sufficient condition is that Z is **uniformly bounded from below**. But this is rather implicit.
  - Z is uniformly bounded from below if there exists an equivalent martingale measure for S which satisfies the **reverse Hölder inequality** $R_2(P)$.
  - There exists an equivalent martingale measure for S which satisfies the reverse Hölder inequality $R_2(P)$ if the **mean–variance tradeoff process** of S is **bounded** and we have $\lambda_k \Delta M_k < 1$ for all k.

Executive summary

Main message:

The **mean–variance portfolio optimisation** problem can be solved with the help of **dynamic programming methods**, in full generality, in finite discrete time.

References

- H. Pham and X. Wei, "Discrete time McKean-Vlasov control problem: a dynamic programming approach", *Applied Mathematics and Optimization* 74 (2016), 487–506
- Z. Tan, "Dynamic Programming Approaches for the Mean–Variance Portfolio Selection Problem", *PhD thesis ETH Zurich, forthcoming* (2022)
- . . . and of course the book:
- T. Björk, M. Khapko and A. Murgoci, "Time-Inconsistent Control Theory with Finance Applications", *Springer* (2021)

Thank you for your attention