Stochastic dynamic optimization with multivariate stochastic dominance constraints

Darinka Dentcheva and **Andrzej Ruszczyński**
Stevens Institute of Technology and Rutgers University

References:

Optimization with stochastic dominance constraints,
SIAM Journal on Optimization 14 (2003) 548–566

Stochastic optimization with nonlinear dominance constraints,
Mathematical Programming 99 (2004) 329–350

Stochastic dynamic optimization with discounted stochastic dominance constraints,
Mathematical Programming, to appear

Composite semi-infinite optimization,
Control and Cybernetics, to appear

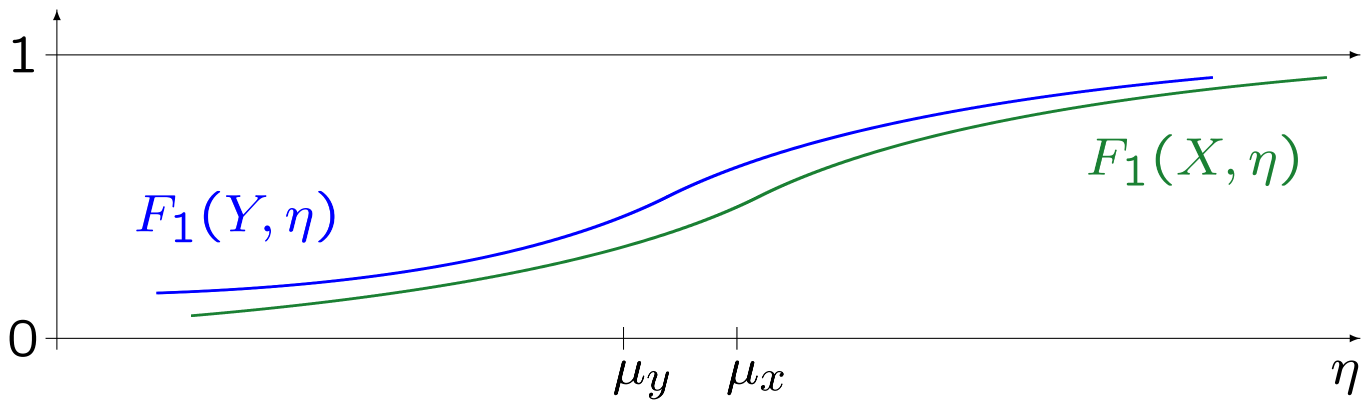
available from `www.rusz.rutgers.edu`

Supported by the National Science Foundation

First Order Dominance

Notation: $F_1(X,\eta) = \mathbb{P}\{X \le \eta\}$

$$X \succeq_{(1)} Y \quad \Leftrightarrow \quad F_1(X,\eta) \le F_1(Y,\eta) \quad \text{for all } \eta \in \mathbb{R}$$



Relation to utility functions:

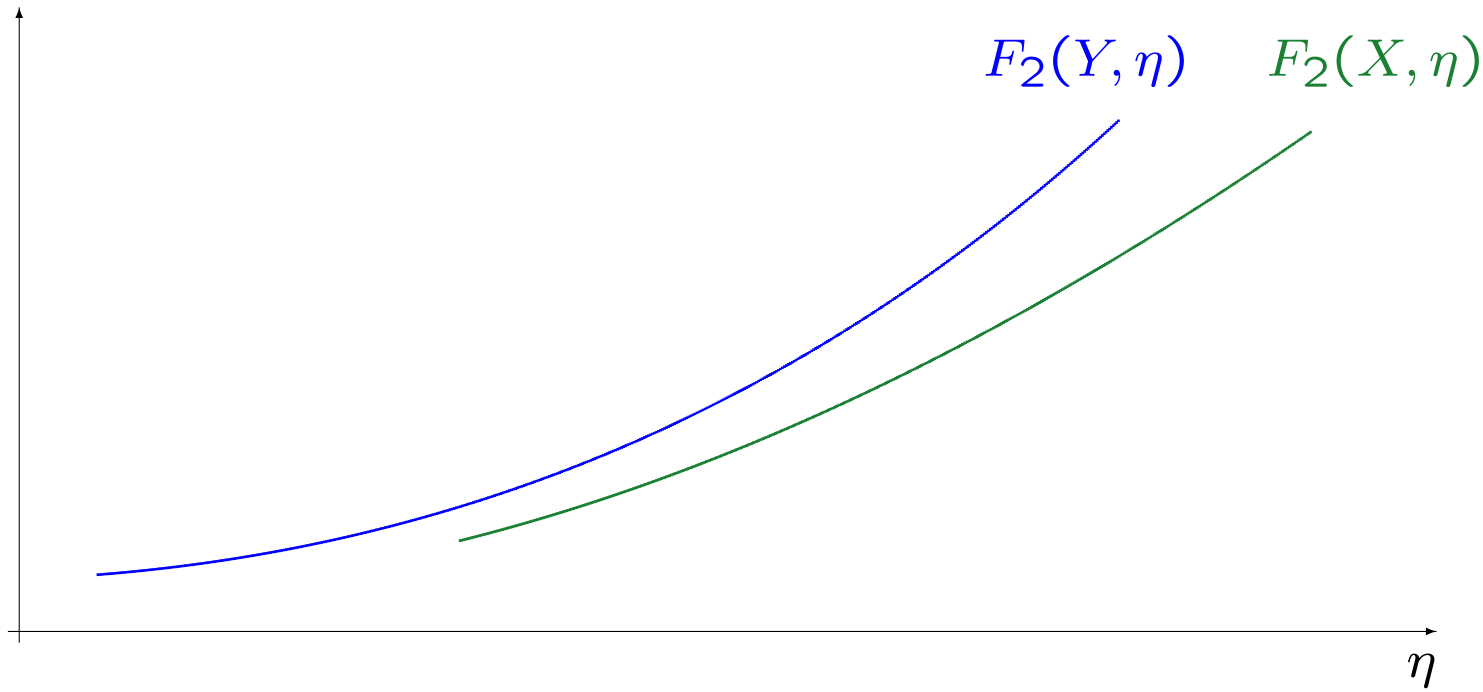
$$X \succeq_{(1)} Y \;\Leftrightarrow\; \mathbb{E}\,u(X) \ge \mathbb{E}\,u(Y) \quad \forall \textit{ nondecreasing } u(\cdot)$$

Quirk and Saposnik (1962)

Second Order Dominance

Notation: $F_2(X,\eta) = \int_{-\infty}^{\eta} F_1(X,\xi)\,d\xi = \mathbb{E}(\eta - X)_+$ for $\eta \in \mathbb{R}$

$$X \succeq_{(2)} Y \quad \Leftrightarrow \quad F_2(X,\eta) \le F_2(Y,\eta) \quad \text{for all } \eta \in \mathbb{R}$$



Relation to utility functions:

$$X \succeq_{(2)} Y \;\Leftrightarrow\; \mathbb{E}\,u(X) \ge \mathbb{E}\,u(Y) \quad \forall \text{ nondecreasing concave } u(\cdot)$$

Hadar and Russell (1969)

Dominance Constrained Optimization Problems

$$\max \ \mathbb{E}\Big[H(z)\Big]$$
$$\text{subject to } G(z) \succeq_{(2)} Y$$
$$z \in Z_0$$

Z_0 - convex closed subset of a Banach space $\mathscr{Z}$
G and H - continuous concave operators from $\mathscr{Z}$ to $\mathscr{L}_1(\Omega, \mathscr{F}, P)$
Y - benchmark outcome in $\mathscr{L}_1(\Omega, \mathscr{F}, P)$

Convenient re-formulation of the dominance constraint:

$$F_2(G(z); \eta) \le F_2(Y; \eta) \quad \text{for all } \eta \in [a, b]$$

How to formulate the order $\succeq_{(2)}$ if the outcome and the benchmark are multivariate?

$$G(z) = \Big(G_1(z), G_2(z), \ldots, G_{T+1}(z)\Big)$$
$$Y = (Y_1, Y_2, \ldots, Y_{T+1})$$

Important case: dynamic with discrete time

Multivariate Stochastic Orders

Consider $X = (X_1, \dots, X_{T+1})$ and $Y = (Y_1, \dots, Y_{T+1})$ in $\mathscr{L}_1^{T+1}(\Omega, \mathscr{F}, P)$.

Increasing Convex Order

$$X \succeq_{(\text{icx})} Y \;\Leftrightarrow\; \mathbb{E}\, u(X) \geq \mathbb{E}\, u(Y) \quad \forall\, u \in \mathscr{U}$$

Generator $\mathscr{U}$ - all concave nondecreasing functions $u : \mathbb{R}^{T+1} \to \mathbb{R}$

Hard to treat analytically, the generator is too rich

Coordinate Order

$$X \succeq_{(2)}^{\text{sep}} Y \;\Leftrightarrow\; X_t \succeq_{(2)} Y_t,\; t = 1, \dots, T+1$$

Generator $\mathscr{U}$ - all functions $u(X) = \sum_{t=1}^{T+1} u_t(X_t)$ with concave nondecreasing $u_t : \mathbb{R} \to \mathbb{R}$. Our earlier analysis covers this case.

Ignores temporal structure and dependency

Our idea is to define the mutlivariate order via a family of univariate orders

Dynamic Stochastic Dominance

Consider $X = (X_1, \ldots, X_{T+1})$ and $Y = (Y_1, \ldots, Y_{T+1})$ in $\mathscr{L}_1^{T+1}(\Omega, \mathscr{F}, P)$.

New idea: Consider linear scalarization by comparing

$$\langle \varrho, X \rangle = \sum_{t=1}^{T+1} \varrho_t X_t \quad \text{and} \quad \langle \varrho, Y \rangle = \sum_{t=1}^{T+1} \varrho_t Y_t$$

for all

$$\varrho \in \mathscr{D} \subseteq \left\{ \varrho \in \mathbb{R}^{T+1} : 1 \geq \varrho_1 \geq \varrho_2 \geq \cdots \geq \varrho_{T+1} \geq 0 \right\}$$

Discounted Order: $X \succeq_{(2)}^{\text{dis}} Y$ if $\langle \varrho, X \rangle \succeq_{(2)} \langle \varrho, Y \rangle$ for all $\varrho \in \mathscr{D}$

The order $\succeq_{(2)}^{\text{dis}}$ neither implies nor is implied by the coordinate order

Special case:

Finite set $\mathscr{D}$ consisting of $T+1$ elements ϱ^k, $k = 1, \ldots, T+1$, with

$$\varrho_t^k = \begin{cases} 1 & \text{if } \; t \le k, \\ 0 & \text{if } \; t > k. \end{cases}$$

Then $X \succeq_{(2)}^{\text{dis}} Y$ is equivalent to the dominance of partial sums

$$\sum_{t=1}^{k} X_t \succeq_{(2)} \sum_{t=1}^{k} Y_t, \quad k = 1, \ldots, T+1.$$

By specifying the set $\mathscr{D}$ we can obtain other interesting cases.

Similar ideas can be employed in the multivariate case, without the dynamic structure

The Generator of $\succeq_{(2)}^{\text{dis}}$

Notation: $\mathscr{M}_+(S)$ - the space of nonnegative regular measures on $S \subset \mathbb{R}^m$

For every $\lambda \in \mathscr{M}_+(\mathbb{R} \times \mathscr{D})$ we define a concave nondecreasing function $\varphi_\lambda : \mathbb{R}^{T+1} \to \mathbb{R}$:

$$\varphi_\lambda(x) = -\int_{\mathbb{R}\times\mathscr{D}} \max(0, \eta - \langle \varrho, x \rangle)\ \lambda(d(\eta, \varrho)).$$

Theorem The class of functions

$$\Phi = \{\varphi_\lambda : \lambda \in \mathscr{M}_+(\mathbb{R} \times \mathscr{D})\}$$

is a generator of the order $\succeq_{(2)}^{\text{dis}}$:

$$X \succeq_{(2)}^{\text{dis}} Y \quad \Leftrightarrow \quad \mathbb{E}[\varphi(X)] \geq \mathbb{E}[\varphi(Y)] \quad \text{for all } \varphi \in \Phi$$

Obtained by integrating w.r.t. marginal and conditional measures of λ

Dynamic Optimization Problem

$$\begin{aligned}
\text{max} \ & \sum_{t=1}^{T} \mathbb{E} G_t(s_t, v_t) + \mathbb{E} G_{T+1}(s_{T+1}) \\
\text{s.t.} \ & s_{t+1} = A_t s_t + B_t v_t + e_t, \quad t = 1, \dots, T, \\
& (G_1(s_1, v_1), \dots, G_1(s_T, v_T), G_{T+1}(s_{T+1})) \succeq_{(2)}^{\text{dis}} (Y_1, \dots, Y_T, Y_{T+1}) \\
& v_t \in V_t \text{ a.s.}, \quad t = 1, \dots, T.
\end{aligned}$$

Decision vector z: controls $v_1, \dots, v_T$, and trajectories $s_2, \dots, s_{T+1}$

$\mathscr{F}_t$ - σ-algebra generated by $\{e_1, \dots, e_{t-1}\}$
Space of controls: $\mathscr{V} = \mathscr{L}_1^{n_v}(\Omega, \mathscr{F}_1, P) \times \cdots \times \mathscr{L}_1^{n_v}(\Omega, \mathscr{F}_T, P)$.
Space of state trajectories: $\mathscr{S} = \mathscr{L}_1^{n_s}(\Omega, \mathscr{F}_2, P) \times \cdots \times \mathscr{L}_1^{n_s}(\Omega, \mathscr{F}_{T+1}, P)$.

For technical reasons we restrict the range of $\eta \in \mathbb{R}$ to an interval $[a, b]$:

$$F_2(\langle \varrho, G(s, v) \rangle; \eta) \le F_2(\langle \varrho, Y \rangle; \eta) \quad \text{for all } \varrho \in \mathscr{D} \text{ and all } \eta \in [a, b]$$

The Partial Lagrangian

Define the class of functions

$$\Phi([a,b],\mathscr{D}) = \{\varphi_\lambda : \lambda \in \mathscr{M}_+([a,b]\times\mathscr{D})\},$$

$$\varphi_\lambda(x) = -\int_{[a,b]\times\mathscr{D}} \max(0, \eta - \langle \varrho, x\rangle)\ \lambda(d(\eta,\varrho))$$

The set $\Phi([a,b],\mathscr{D})$ is a generator of the interval dominance order

We introduce the partial "Lagrangian" $L : \mathscr{S} \times \mathscr{V} \times \Phi([a,b],\mathscr{D}) \to \mathbb{R}$:

$$L(s,v,\varphi) = \mathbb{E}\Big[\sum_{t=1}^{T} G_t(s_t,v_t) + G_{T+1}(s_{T+1})$$
$$+ \Big(\varphi(G_1(s_1,v_1),\dots,G_T(s_T,v_T),G_{T+1}(s_{T+1})) - \varphi(Y_1,\dots,Y_T,Y_{T+1})\Big)\Big]$$

Non-decomposable w.r.t. $t = 1,\dots,T+1$

Auxiliary Control Problem

$$\begin{aligned}
\max \ \mathbb{E}\Big[\sum_{t=1}^{T} G_t(s_t, v_t) + G_{T+1}(s_{T+1}) \\
+ \Big(\varphi(G_1(s_1, v_1), \dots, G_T(s_T, v_T), G_{T+1}(s_{T+1})) - \varphi(Y_1, \dots, Y_T, Y_{T+1}) \Big) \Big]
\end{aligned}$$

$$\begin{aligned}
\text{s.t.} \quad & s_{t+1} = A_t s_t + B_t v_t + e_t, \quad t = 1, \dots, T \\
& v_t \in V_t \quad \text{a.s.}, \quad t = 1, \dots, T
\end{aligned}$$

Uniform Dominance Condition: $\exists$ a feasible pair $(\tilde{s}, \tilde{v})$ such that

$$\inf_{(\eta,\varrho)\in[a,b]\times\mathscr{D}} \Big\{ F_2(\langle \varrho, Y \rangle; \eta) - F_2(\langle \varrho, G(\tilde{s}, \tilde{v}) \rangle; \eta) \Big\} > 0.$$

Theorem

Assume the uniform dominance condition. If $(\hat{s}, \hat{v})$ is an optimal solution of the original problem then there exist $\hat{\varphi} \in \Phi([a, b], \mathscr{D})$ such that $(\hat{s}, \hat{v})$ is an optimal solution of the auxiliary problem with $\varphi = \hat{\varphi}$, and

$$\mathbb{E}[\hat{\varphi}(G(\hat{s}, \hat{v}))] = \mathbb{E}[\hat{\varphi}(Y)].$$

Conversely, if for some $\hat{\varphi} \in \Phi([a, b], \mathscr{D})$ an optimal solution $(\hat{s}, \hat{v})$ of the auxiliary problem satisfies the dominance constraint, then it is optimal.

Ideas of the Derivation

1. Write the dominance constraint as an operator constraint in $\mathscr{C}([a,b],\mathscr{D})$:

$$\mathbb{E}(\eta - \langle \varrho, G(s,v)\rangle)_+ \le \mathbb{E}(\eta - \langle \varrho, Y\rangle)_+ \quad \text{for all } \varrho \in \mathscr{D} \text{ and all } \eta \in [a,b]$$

2. Associate with it a Lagrange multiplier

   - nonnegative measure $\lambda \in \mathscr{M}_+([a,b]\times\mathscr{D})$

3. Change the order of integration in the Lagrangian term

$$\int_{[a,b]\times\mathscr{D}} \mathbb{E}(\eta - \langle \varrho, X\rangle)_+ \, \lambda(d(\eta,\varrho)) = \int_\Omega \int_{[a,b]\times\mathscr{D}} (\eta - \langle \varrho, X(\omega)\rangle)_+ \, \lambda(d(\eta,\varrho)) \, P(d\omega)$$
$$= -\mathbb{E}\varphi(X)$$

with

$$\varphi(x) = -\int_{[a,b]\times\mathscr{D}} (\eta - \langle \varrho, x\rangle)_+ \, \lambda(d(\eta,\varrho))$$

Implied Random Discount

$$\begin{aligned}
\max \quad & \sum_{t=1}^{T} \mathbb{E}G_t(s_t, v_t) + \mathbb{E}G_{T+1}(s_{T+1}) \\
\text{s.t.} \quad & s_{t+1} = A_t s_t + B_t v_t + e_t, \quad t = 1, \dots, T, \\
& (G_1(s_1, v_1), \dots, G_1(s_T, v_T), G_{T+1}(s_{T+1})) \succeq_{(2)}^{\text{dis}} (Y_1, \dots, Y_T, Y_{T+1}) \\
& v_t \in V_t \text{ a.s.}, \quad t = 1, \dots, T.
\end{aligned}$$

Theorem Assume the uniform dominance condition. If $(\hat{s}, \hat{v})$ is an optimal solution of the original control problem then there exist $\xi_t \in \mathscr{L}_\infty(\Omega, \mathscr{F}_t, P)$, $t = 1, \dots, T+1$, with

$$\xi_1 \geq \xi_2 \geq \cdots \geq \xi_T \geq \xi_{T+1} \geq 0 \quad \text{a.s.}$$

such that $(\hat{s}, \hat{v})$ is an optimal solution of the control problem

$$\begin{aligned}
\max \quad & \sum_{t=1}^{T} \mathbb{E}(1 + \xi_t) G_t(s_t, v_t) + \mathbb{E}(1 + \xi_{T+1}) G_{T+1}(s_{T+1}) \\
\text{s.t.} \quad & s_{t+1} = A_t s_t + B_t v_t + e_t, \quad t = 1, \dots, T, \\
& v_t \in V_t \quad \text{a.s.}, \quad t = 1, \dots, T.
\end{aligned}$$

Maximum Principle

Theorem Assume the uniform dominance condition. If $(\hat{s}, \hat{v})$ is an optimal solution of the original control problem, then there exist:
discount factors $\xi_t \in \mathscr{L}_\infty(\Omega, \mathscr{F}_t, P)$, $t = 1, \ldots, T+1$,

$$\xi_1 \geq \xi_2 \geq \cdots \geq \xi_T \geq \xi_{T+1} \geq 0 \quad \text{a.s.}$$

subgradients $(\sigma_t^s, \sigma_t^v) \in \mathscr{L}_q^{n_s}(\Omega, \mathscr{F}_t, P) \times \mathscr{L}_q^{n_v}(\Omega, \mathscr{F}_t, P)$ satisfying

$$(\sigma_t^s(\omega), \sigma_t^v(\omega)) \in \partial g_t(\hat{s}_t(\omega), \hat{v}_t(\omega))$$
$$\sigma_{T+1}^s(\omega) \in \partial g_{T+1}(\hat{s}_{T+1}(\omega))$$

and dual variables $\hat{y}_t \in \mathscr{L}_q^{n_s}(\Omega, \mathscr{F}_{t+1}, P)$, $t = 1, \ldots, T$, satisfying the adjoint equations

$$y_T = (1 + \xi_{T+1})\sigma_{T+1}^s$$
$$y_{t-1} = A_t' \mathbb{E}[y_t | \mathscr{F}_t] + (1 + \xi_t)\sigma_t^s, \quad t = T, \ldots, 2$$

such that for almost all $\omega \in \Omega$ the control $\hat{v}_t(\omega)$ is a solution of

$$\max_{c \in V_t} \left\langle (1 + \xi_t(\omega))\sigma_t^v(\omega) + B_t'(\omega)\mathbb{E}[y_t | \mathscr{F}_t](\omega), c \right\rangle$$