On portfolio optimization under incomplete information

Wolfgang J. Runggaldier,
Dipartimento di Matematica Pura ed Applicata
University of Padova, Italy

*Workshop Stochastic Analysis and Applications in Finance.
Max Planck Institute for Mathematics in the Sciences,
In general, for portfolio optimization two main solution approaches:

i) *Dynamic Programming (DP)*

ii) “Martingale method” (convex duality)

Here essentially only the DP-based approach (due to its intrinsic dynamic nature it is better suited for information that grows with time)

→ Convex duality so far only for special cases of incomplete information (Lakner 95/98, Karatzas-Zhao 01, Zohar 01, Haussmann-Sass 04).

→ With respect to convex duality, for DP it is irrelevant whether the market is complete or not (incomplete models in the full filtration may become complete in the observed sub-filtration).
Outline

I. Discrete time

- Complete information models (financial context)

- DP for complete information (DP principle and DP backwards recursions)

- Incomplete information models (financial context)

- Reformulation as complete information problem (discrete time filtering equations)

- Discussion of computational aspects (approximations; certainty equivalence)
II. Continuous time

- Complete information models
  
- DP for complete information: HJB equations \((\textit{heuristic derivation from discrete time recursions; verification theorem})\)

- Incomplete information models
  
- Reformulation as complete information problem \((\textit{General nonlinear filtering equations; finite-dimensional filters})\)

- Example: the case of a power utility \((\textit{reformulation as a complete information “risk sensitive problem” and its DP-solution})\).

III. Jump process models

- A special class of models \((\textit{certainty equivalence for log and power utility; solution approaches})\).

→ no transaction costs
I. Discrete time

A. Complete information

Financial context (portfolio optimization)

- A market model with $B_t \equiv 1$ and $S_t$ s.t.
  \[ dS_t = S_t \left[ a \, dt + \sigma \, dw_t \right] \]

- If investment decisions at discrete time intervals (step $\Delta > 0$), need $S_t$ only at discrete time points

→ Standard Ito formula leads to
  \[ S_{t+1} = S_t \exp \left[ \left( a - \frac{\sigma^2}{2} \right) \Delta + \sigma \Delta w_t \right] := S_t \xi_t \]

→ Here $\xi_t$ i.i.d. lognormal. Putting $Z_t := \log S_t$, 
  \[ Z_{t+1} = Z_t + \log \xi_t \]

→ $\xi_t$ could also be i.i.d. multinomial:
  \[ \xi_t \in \{\xi^1, \ldots, \xi^K\} \text{ with prob. } q = (q^1, \ldots, q^K) \]
Letting
\[ \alpha_t : \text{no. units of } S_t; \]
\[ c_t : \text{consumption}; \]
\[ V_t : \text{portfolio value}, \]
the self financing condition requires
\[ \Delta V_t = \alpha_t \Delta S_t - c_t = \alpha_t S_t(\xi_t - 1) - c_t \]

→ Putting \( h_t := \frac{\alpha_t S_t}{V_t} \) (ratio process):
\[ V_{t+1} = V_t + V_t h_t(\xi_t - 1) - c_t \]
\[ = [V_t(1 - h_t) - c_t] + V_t h_t \xi_t \]

Criteria :

i) **Maximization of expected utility from consumption and terminal wealth**
\[ \sup_{h,c} E \left\{ \sum_{t=0}^{T-1} U_t(c_t) + U_T(V_T - c_T) \right\} \]

ii) **Hedging of a contingent claim**
\[ \sup_h E \left\{ -U_T \left( [H_T(e^{Z_t}) - V_T]^+ \right) \right\} \]
General complete information model

- Put $Y_t := [Z_t, V_t]'$, $H_t := [h_t, c_t]$, $D_t := \xi_t,$

\[
\begin{aligned}
Y_{t+1} &= G_t(Y_t, H_t, D_t) , \quad t = 0, \ldots, T \\
&\quad \sup_{H} \mathbb{E} \left\{ \sum_{t=0}^{T} U_t(Y_t, H_t) \right\}
\end{aligned}
\]

→ distribution of $D_t$ ("disturbances") may depend on $Y_t$, but not on its past

→ $H_t \in \mathcal{F}^{Y}_t$ ("full filtration")

→ if $H_t = H_t(Y_t)$ "Markov controls" then $Y_t$ is Markov

→ for the given model $\exists$ a Markov optimal control.
B. Dynamic Programming

Dynamic Programming Principle

- If a process is optimal over an entire sequence of periods, then it has to be optimal over each single period.

→ Allows to determine an optimal sequence $h_0, \ldots, h_T$ by a sequence of minimizations over the individual controls.

- Due to the Markovianity of $H_t$, and thus of $Y_t$, and the additivity over time of the criterion

$$
\sup_{H_0, \ldots, H_T} E \left\{ \sum_{t=0}^{T} U_t(Y_t, H_t) \right\}
$$

$$
= \sup_{H_0} [U_0(Y_0, H_0) + E\left\{ \sup_{H_1} [U_1(Y_1, H_1) + E\{ \ldots \ldots \ldots \ldots \ldots \ldots \}
\right. 
\left. \ldots \ldots \ldots \ldots \ldots \ldots \right. 
\left. + E\left\{ \sup_{H_T} U_T(Y_T, H_T) \mid (Y, H)_{T-1} \right\} \right\} \mid (Y, H_0) ]
$$
DP-algorithm

- Recursive implementation of the DP principle: let

\[
\Phi_t(y) := \sup_{H_t, \ldots, H_T} E \left\{ \sum_{s=t}^T U_s(Y_s, H_s) \mid Y_t = y \right\}
\]

then (DP-algorithm)

\[
\begin{align*}
\Phi_T(y) &= \sup_H U_T(y, H) \\
\Phi_t(y) &= \sup_H [U_t(y, H) + E\{\Phi_{t+1}(G_t(y, H, D_t)) \mid Y_t = y\}]
\end{align*}
\]

→ The recursive backwards induction requires, in each period \( t \) and for each possible value \( Y_t = y \) of \( Y \) in that period, a maximization over a single \( H \) which, at the same time, leads to an optimal Markov strategy.
**Difficulty**: the growing complexity with increasing $t$ ("curse of dimensionality") due to the growing number of possible values of $y$ (all values of $\Phi(y)$ have to be stored for each $t$)

$\rightarrow$ *Explicit analytical solutions are possible in particular cases.*

- "Linear-quadratic Gaussian (LQG)" : dynamics linear, disturbances Gaussian, utility functions (negative) quadratic.

  $\rightarrow$ $H_t^*(Y_t) = K_t \cdot Y_t$ ; $\Phi_t(y)$ quadratic

- Optimal control belongs to a finitely parametrized family of Markov controls (optimization over a finite-dimensional parameter)

  $\rightarrow$ *In general : approximation methods (quantization)*
Incomplete information

Financial context

- Start from the continuous-time price process model

\[
\begin{align*}
    dS_t &= S_t [a_t(X_t)dt + \sigma_t(X_t)dw_t] \\
    X_t &\quad \text{hidden factor process} \\
    &\quad \text{(diffusion or finite-state Markov)}
\end{align*}
\]

\( \rightarrow \) after integration (step \( \Delta > 0 \))

\[
\begin{align*}
    S_{t+1} &= S_t \exp \left[ \left( a_t(X_t) - \frac{\sigma_t^2(X_t)}{2} \right) \Delta + \sigma_t(X_t)\Delta w_t \right] \\
    &\quad := S_t \bar{\xi}_t \\
    X_{t+1} &= (\Lambda_t X_t + \Gamma_t) + M_t \\
    M_t &\quad : \text{martingale difference sequence}
\end{align*}
\]

\( \rightarrow \) as before

\[
V_{t+1} = [V_t(1 - h_t) - c_t] + V_th_t \bar{\xi}_t
\]

\( \rightarrow \) Distribution of \( \bar{\xi}_t \) now depends explicitly on \( X_t \).
General model

- Putting, as before, $Z_t := \log S_t$, $Y_t := [Z_t, V_t]'$, $H_t := [h_t, c_t]'$, $D_t = \bar{\xi}_t$

\[
\begin{align*}
X_{t+1} &= (\Lambda_t X_t + \Gamma_t) + M_t \quad \text{(unobserved)} \\
Y_{t+1} &= G_t(Y_t, X_t, H_t, D_t) \quad \text{(observed)} \\
\sup_H E \left\{ \sum_{t=0}^{T} U_t(Y_t, X_t, H_t) \right\}
\end{align*}
\]

thereby including, for full generality, $X_t$ in the dynamics of the criterion ($\Lambda_t, \Gamma_t$ could depend also on $Y_t$)

$\rightarrow (X_t, Y_t)$ is Markov

$\rightarrow H_t \in \mathcal{F}_{t}^{Y}$ but here the “full filtration” is $\mathcal{F}_{t}^{X,Y}$ with $\mathcal{F}_{t}^{Y} \subset \mathcal{F}_{t}^{X,Y}$

$\rightarrow$ Unknown parameters may be incorporated as constant components of $X_t$ (in the multinomial model, if $q = (q^1, \cdots, q^K)$ are not fully known, they become unknown constant parameters).
Reformulation as complete information problem ("separated problem")

- Suppose that, independently of the control policy $H$, $\exists$ an observable controlled Markov process $\{\Psi_t\}$ s.t., for a suitable $u_t(\cdot)$, one has

$$u_t(\Psi_t, H_t) = E\left\{U_t(Y_t, X_t, H_t) \mid (Y^t_0, H^t_0)\right\}$$

$\Psi_t$ is a "sufficient statistic" synthesizing the information on $X_t$ obtained from the current history of observations and controls that is relevant for the control purposes.

→ One then has in fact

$$E\left\{\sum_{t=0}^T U_t(Y_t, X_t, H_t)\right\} = \sum_{t=0}^T E\left\{E\left\{U_t(Y_t, X_t, H_t) \mid (Y^t_0, H^t_0)\right\}\right\}$$

$$= \sum_{t=0}^T E\left\{u_t(\Psi_t, H_t)\right\} = E\left\{\sum_{t=0}^T u_t(\Psi_t, H_t)\right\}$$

→ $(\Psi_t, V_t)$ may be taken to play the role of $Y_t$ in the complete information model and $\exists$ a Markov-optimal control $H_t = H_t(\Psi_t, V_t)$. 
Typical case: conditional (filter) distribution

Take \( \Psi_t := [p(X_t \mid Z_t^0), Y_t] \)

→ Since, for any given strategy \( H \), \( \mathcal{F}_t^Y = \mathcal{F}_t^Z \), we have \( p(X_t \mid (Y_t^0, H_t^0)) = p(X_t \mid Z_t^0) \).

→ \( \Psi_t \) can be shown to be a Markov process (controlled because of the \( Y \)-component).

→ By the recursive Bayes formula (discrete time filter equation)

\[
p(X_{t+1} \mid Z_{0}^{t+1}) \propto p(X_{t+1}, Z_{t+1} \mid Z_{0}^{t})
= \int p(X_{t+1}, X_t, Z_{t+1} \mid Z_{0}^{t}) d\mu(X_t)
= \int p(Z_{t+1} \mid Z_t, X_t) \cdot p(X_{t+1} \mid X_t, Z_{0}^{t}) d\mu(X_t)
= \int p(Z_{t+1} \mid Z_t, X_t) \cdot p(X_{t+1} \mid X_t, Z_t) dp(X_t \mid Z_{0}^{t})
\]

→ \( Z_t \) plays the role of a disturbance (its distribution depends on the current state).
• One also has
\[ E \left\{ U_t(Y_t, X_t, H_t) \mid (Y_0^t, H_0^t) \right\} \]
\[ = \int U_t(Y_t, x, H_t)dp(x \mid (Y_0^t, H_0^t)) = \int U_t(Y_t, x, H_t)dp(x \mid Z_0^t) \]
\[ := u_t(\Psi_t, H_t). \]

• One may consider an un-normalized version of \( p(X_t \mid Z_0^t) \). It satisfies a linear recursion ("Zakai equation"). The criterion has then to be expressed as expectation under \( Q \sim P \) ("reference probability measure") under which \( Y_t \) becomes independent of \( X_t \).

→ We have obtained an equivalent complete information problem with new “state” \( \Psi_t = [p(X_t \mid Z_0^t), Y_t] \) that in general is \( \infty \)–dimensional or, at least, \( \infty \)–valued (even if \( X_t \) is finite-valued, \( p(X_t \mid Z_0^t) \) takes values in a simplex).
• There are cases when \( p(X_t \mid Z^t_0) \) is finitely parametrized and then \( \Psi_t \) is finite-dimensional.

• Alternatively, there are cases where the optimal control belongs to a finitely parametrized family of Markov controls (optimization over a finite-dimensional parameter).

• Otherwise approximation (quantization).

• In the case of an unknown parameter \( \theta \), one may put \( X_t \equiv \theta \) and the recursive Bayes formula for \( p(X_t \mid Z^t_0) \) reduces then to the ordinary Bayes formula for \( p(\theta \mid Z^t_0) \).

\[
\rightarrow \text{E.g. the multinomial model with } \theta = q, \text{namely the vector of (unknown) probabilities. Here } p(q \mid Z^t_0) \text{ may be taken in the finitely parametrized class of Dirichlet distributions.}
\]
• **Certainty-equivalence property (CE)**: the optimal control for incomplete information is the same as for complete information with the unknown random quantities replaced by their conditional means.

→ *CE can hold only in cases where the control has no “dual effect”: it does not enhance further learning to improve overall performance.*

**Examples**

• Linear-quadratic Gaussian models (*standard control problem*).

If under complete information

\[ H^*_t(X_t) = K_tX_t \]

then, under incomplete information,

\[ H^*_t(\Psi_t) = K_tE_{\Psi_t}(X_t) \]

• Hedging with a quadratic penalization of the hedging error
• Multinomial models with incomplete information on \( q = (q^1, \cdots, q^K) \)

→ DP under complete information is

\[
\begin{align*}
\Phi_T(y) &= \sup_H U_T(y, H) \\
\Phi_t(y) &= \sup_H \left[ U_t(y, H) + \sum_{k=1}^{K} q^k \Phi_{t+1}(G_t(y, H, \xi^k)) \right]
\end{align*}
\]

→ Under incomplete information, by the DP principle and the fact that here \( U_t(Y_t, X_t, H_t) \) does not depend on \( X_t \equiv q \), the calculations lead to

\[
\sup_H \left[ U_t(y, H) + \mathbb{E} \left\{ \sum_{k=1}^{K} q^k \Phi_{t+1}(G_t(y, H, \xi^k)) \mid (Y^0_t, H^0_t) \right\} \right]
\]

\[
= \sup_H \left[ U_t(y, H) + \sum_{k=1}^{K} \mathbb{E} \left\{ q^k \mid Z^t_0 \right\} \Phi_{t+1}(G_t(y, H, \xi^k)) \right]
\]

→ The DP algorithm is the same as in the complete information case except that \( q \) is replaced in the generic period \( t \) by \( \mathbb{E} \left\{ q \mid Z^t_0 \right\} \).
II. Continuous time

A. Complete information

Financial context and model

\[
\begin{align*}
    dS_t &= S_t [a \, dt + \sigma \, dw_t] ; \quad B_t \equiv 1 \\
    dV_t &= \alpha_t S_t - c_t \, dt
\end{align*}
\]

\[ \rightarrow \text{with } Z_t := \log S_t \text{ and } h_t := \frac{\alpha_t S_t}{V_t} \]

\[
\begin{align*}
    dZ_t &= \left( a - \frac{\sigma^2}{2} \right) dt + \sigma \, dw_t \\
    dV_t &= (V_t a h_t - c_t) dt + V_t \sigma h_t \, dw_t
\end{align*}
\]

\[ \rightarrow \text{with } Y_t := [Z_t, V_t]' ; \quad H_t := [h_t, c_t] \text{ it leads to a control problem of the following general form} \]

\[
\begin{align*}
    dY_t &= A_t(Y_t, H_t) dt + B_t(Y_t, H_t) dw_t ; \quad t \in [0, T] \\
    \sup_H E \left\{ \int_0^T U_t(Y_t, H_t) dt + U_T(Y_T, H_T) \right\}
\end{align*}
\]

where \( H_t \in \mathcal{F}_t^Y \).
B. Dynamic Programming (HJB equation)

Heuristic derivation (from discrete to continuous time)

- Apply an Euler-type discretization (step $\Delta$)

\[
\begin{align*}
Y_{t+1} &= Y_t + A_t(Y_t, H_t) \Delta + B_t(Y_t, H_t) \Delta w_t \\
&:= G_t(Y_t, H_t, \Delta w_t) \\
\sup_{H} E \left \{ \Delta \sum_{t=0}^{T/\Delta} U_t(Y_t, H_t) \right \}
\end{align*}
\]

and recall that

\[
\Phi_t(y) = \sup_{H} [\Delta U_t(y, H) + E \{ \Phi_{t+1}(G_t(y, H, \Delta w_t)) | Y_t = y \}]
\]

Via a Taylor expansion and $E(\Delta w_t)^2 \approx \Delta$

\[
\sup_{H} \left [ \Delta U_t(y, H) + E \{ \Phi_{t+1} - \Phi_t(y) | Y_t = y \} \right ]
\]

\[
= \frac{\partial}{\partial t} \Phi_t(y) \Delta + \sup_{H} \left [ \Delta U_t(y, H) + \Delta A_t(y, H) \frac{\partial}{\partial y} \Phi_t(y) \\
+ \frac{\Delta}{2} B_t^2(y, H) \frac{\partial^2}{\partial y^2} \Phi_t(y) + o(\Delta) \right ] = 0
\]
Dividing by $\Delta$ and letting $\Delta \downarrow 0$, 

\[
\begin{align*}
\frac{\partial}{\partial t} \Phi_t(y) + \sup_H \left[ A_t(y, H) \frac{\partial}{\partial y} \Phi_t(y) \\
+ \frac{1}{2} B_t^2(y, H) \frac{\partial^2}{\partial y^2} \Phi_t(y) + U_t(y, H) \right] &= 0 \\
\Phi_T(y) &= \sup_H U_T(y, H)
\end{align*}
\]

→ **Standard heuristic derivation** (based on the DP principle)

\[
\Phi_t(y) = \sup_{H_s, s \in [t, t+\Delta]} E \left\{ \int_t^{t+\Delta} U_s(Y_s^H, H_s) ds \\
+ \Phi_{t+\Delta}^H(Y_{t+\Delta}^H) \mid Y_t = y \right\}
\]

and then proceed analogously as before.
Solution procedure

i) Solve the maximization over $H$ depending on the yet unknown $\Phi_t(y)$.

ii) Insert the maximizing value $H^*(t, y)$ and solve the resulting PDE.

→ A “verification theorem” guarantees, under sufficient regularity (classical solution), the optimality of the resulting $H^*(t, y)$ and $\Phi_t(y)$.

→ In the absence of sufficient regularity: viscosity solution.

→ As in discrete time, explicit analytical solutions only in particular cases (e.g. linear-quadratic Gaussian).
C. Incomplete information
(Utility from terminal wealth; no consumption)

• Given is the financial model

\[
\begin{align*}
    dS_t &= S_t [a_t(X_t)dt + \sigma_t dw_t] \\
    dX_t &= F_t(X_t)dt + R_t(X_t) dM_t \\
    dV_t &= V_t [h_t a_t(X_t)dt + h_t \sigma_t dw_t]
\end{align*}
\]

→ \(M_t\) a martingale independent of \(w_t\)

→ \(\sigma_t\) is independent of \(X_t\): in continuous time \(\int_0^t \sigma_s^2 ds\) can be estimated by the empirical quadratic variation (dependence on \(X_t\) ⇒ filter degenerates)

• Putting \(Z_t := \log S_t\), consider the (specific) problem

\[
\begin{align*}
    dX_t &= F_t(X_t)dt + R_t(X_t) dM_t \quad \text{(unobserved)} \\
    dZ_t &= A_t(Z_t, X_t)dt + B(Z_t) dw_t \quad \text{(observed)} \\
    dV_t &= V_t \left[ H_t \left( A_t(Z_t, X_t) + \frac{1}{2} B_t^2(Z_t) \right) dt + H_t B_t(Z_t) dw_t \right] \\
    \sup_H E \left\{ V_T^\mu \right\}, \quad \mu \in (0, 1)
\end{align*}
\]
D. Reformulation of the incomplete information problem ("separated problem")

• Take as new "state"

\[ \Psi_t = p_t(x) = p(X_t | \mathcal{F}_t^Z) \mid X_t=x \]

→ filter distribution of \( X_t \) given \( \mathcal{F}_t^Z \)

• Denote by \( \mathcal{L} \) the infinitesimal generator of \( X_t \) (Markov)

• For \( \phi_t = \phi_t(X_t) \) let

\[ \pi_t(\phi) := E \left\{ \phi(X_t) \mid \mathcal{F}_t^Z \right\} = \int \phi(x) dp_t(x) \]
Putting

\[ A_t(Z_t, p_t) := \pi_t(A_t) = \int A_t(Z_t, x)p_t(x) \]

define the “innovations process”

\[ d\bar{w}_t := B_t^{-1}(Z_t) [dZ_t - \pi_t(A_t)dt] \]

→ It implies a translation of the \((P, \mathcal{F}_t)\)–Wiener \(w_t\):

\[ d\bar{w}_t = dw_t + B_t^{-1}(Z_t) [A_t(Z_t, X_t) - A_t(Z_t, p_t)] dt \]

and thus the implicit measure transformation \(P \rightarrow \bar{P}\) with

\[
\frac{d\bar{P}}{dP}|_{\mathcal{F}_T} = \exp \left\{ \int_0^T [A_t(Z_t, p_t) - A_t(Z_t, X_t)] B_t^{-1}(Z_t)dw_t \right.
\]
\[
- \frac{1}{2} \int_0^T [A_t(Z_t, p_t) - A_t(Z_t, X_t)]^2 B_t^{-2}(Z_t) dt \}
\]
General nonlinear filtering equation
("innovations approach")

\[ \pi_t(\phi) = \pi_0(\phi) + \int_0^t \pi_s(\mathcal{L}\phi) ds \]

\[ + \int_0^t [\pi_s(A_s\phi) - \pi_s(A)\pi_s(\phi)] \, d\bar{w}_s \]

→ If \( p_t(x) \) has a density \( p(t, x) \) then, under regularity and by integration by parts,

\[ dp(t, x) = \mathcal{L}^* p(t, x) dt + p(t, x) [A_t - \pi_t(A_t)] \, d\bar{w}_t \]

→ Difficulty not only due to \( \infty - \)dimensionality (stochastic PDE), but also to the term

\[ \pi_t(A_t) = \int A_t(Z_t, x) p(t, x) dx \]

→ For unnormalized densities a linear stochastic PDE ("Zakai equation")
• DP (HJB-eqn.) difficult for an $\infty-$dimensional state $\Psi_t$

$\rightarrow$ In some cases $p_t(x) = p(X_t \mid F^Z_t) \mid X_t = x$ is finitely parametrized

Examples

• Linear-Gaussian models

$$\begin{cases} dX_t = F_t X_t dt + R_t dv_t \\ dZ_t = A_t X_t dt + B_t dw_t \end{cases}$$

$\{v_t\}, \{w_t\}$ independent standard Wiener.

$\rightarrow$ $p(X_t \mid F^Z_t) \mid X_t = x \sim \mathcal{N}(m_t, \gamma_t)$

$$\begin{cases} dm_t = F_t m_t dt + \gamma_t \frac{A_t}{B_t} d\bar{w}_t \\ \dot{\gamma}_t = 2F_t \gamma_t - \gamma_t^2 \left(\frac{A_t}{B_t}\right)^2 + R_t^2 \end{cases}$$

with $d\bar{w}_t := B_t^{-1}[dZ_t - A_t m_t dt]$
• $X_t$: finite-state Markov with states \{s_1, \cdots, s_K\}, i.e.

$$dX_t = \Lambda_t X_t dt + dM_t$$

→ Put $\phi_i(x) = \begin{cases} 1 & \text{if } x = s_i \\ 0 & \text{if } x \neq s_i \end{cases}$ and let

$$p_t^i := P\left\{X_t = s_i \mid \mathcal{F}_t^Z\right\}$$

Furthermore, put

$$D_t(Z_t) := \text{diag}(A_t(Z_t, s_1), \cdots, A_t(Z_t, s_K))$$

$$A_t(Z_t) := [A_t(Z_t, s_1), \cdots, A_t(Z_t, s_K)]'$$

then, with $p_t = [p_1^t, \cdots, p_K^t]'$,

$$dp_t = \Lambda_t p_t dt + [D_t(Z_t) - (A_t'(Z_t) p_t) I]$$

$$\cdot p_t B_t^{-2}(Z_t)(dZ_t - A_t'(Z_t) p_t dt)$$

→ $p_t$ evolves on a finite-dimensional simplex.
Reformulation in case of a general finite-dimensional filter

• Let

\[ p_t(x) = p(X_t | \mathcal{F}_t^Z)|_{X_t=x} = p(x; \xi_t); \quad \xi_t \in \mathbb{R}^p \]

and suppose that

\[ d\xi_t = \beta_t(Z_t, \xi_t)dt + \delta_t(Z_t, \xi_t)\, d\bar{w}_t \]

• Putting

\[ A_t(Z_t, \xi_t) := \int A_t(Z_t, x)dp(x; \xi_t) \]

on \((\Omega, \mathcal{F}, \mathcal{F}_t, \bar{P})\) with Wiener \(\bar{w}_t\):

\[
\begin{align*}
\frac{d\xi_t}{dt} &= \beta_t(Z_t, \xi_t)dt + \delta_t(Z_t, \xi_t)\, d\bar{w}_t \\
\frac{dZ_t}{dt} &= A_t(Z_t, \xi_t)dt + B_t(Z_t)\, d\bar{w}_t \\
\frac{dV_t}{dt} &= V_t \left[ H_t \left( A_t(Z_t, \xi_t) + \frac{1}{2} B_t^2(Z_t) \right) dt + H_t B_t(Z_t)\, d\bar{w}_t \right] \\
\sup_H \bar{E} \left\{ V_T^\mu \right\} , \quad \mu \in (0, 1)
\end{align*}
\]

→ It is the “separated problem” (equivalent complete information problem)

→ With \(Y_t := [Z_t, \xi_t, V_t]\) it is of the form of the general complete information problem.
Reformulation as a “risk-sensitive” problem
(Nagai-R. 05)

• By Itô’s formula and putting

\[ \eta_t(z, \xi, H) := \frac{1-\mu}{2} H^2 B_t^2(z) - H \left( A_t(z, \xi) + \frac{1}{2} B_t^2(z) \right) \]

one has \( (V_0 = v) \)

\[ dV_t^\mu = V_t^\mu \left\{ -\mu \eta_t(Z_t, \xi_t, H_t)dt + \mu H_t B_t(Z_t) d\tilde{w}_t \right\} \]

\[ \rightarrow \quad V_T^\mu = v^\mu \exp \left\{ -\mu \int_0^T \eta_t(Z_t, \xi_t, H_t)dt \right. \]

\[ + \mu \int_0^T H_t B_t(Z_t) d\tilde{w}_t - \frac{\mu^2}{2} \int_0^T H_t^2 B_t^2(Z_t)dt \left. \right\} \]

\[ \rightarrow \quad A \text{ natural measure transformation is now} \]

\[ \frac{d\tilde{P}}{dP} = \exp \left\{ \mu \int_0^T H_t B_t(Z_t) d\tilde{w}_t - \frac{\mu^2}{2} \int_0^T H_t^2 B_t^2(Z_t)dt \right\} \]

implying that \( \tilde{w}_t \) with

\[ d\tilde{w}_t = d\bar{w}_t - \mu H_t B_t(Z_t)dt \]

is a \( (\tilde{P}, \mathcal{F}_t) – \text{Wiener.} \)
• On \((\Omega, \mathcal{F}, \mathcal{F}_t, \bar{P})\) with Wiener \(\tilde{w}_t\)

\[
\begin{align*}
\frac{d\xi_t}{dt} &= [\beta_t(Z_t, \xi_t) + \mu H_t B_t(Z_t) \delta_t(Z_t, \xi_t)]dt + \delta_t(Z_t, \xi_t)d\tilde{w}_t \\
\frac{dZ_t}{dt} &= [A_t(Z_t, \xi_t) + \mu H_t B_t^2(Z_t)]dt + B_t(Z_t)d\tilde{w}_t \\
\sup_H \bar{E}\{V_T^{\mu}\} &= v^\mu \sup_H \bar{E}\left\{\exp \left[ -\mu \int_0^T \eta_t(Z_t, \xi_t, H_t) dt \right] \right\}
\end{align*}
\]

→ No more need for explicit dynamics of \(V_t\) (need only the value of \(V_0 = v\)).

→ A complete observation problem of the "risk-sensitive" type for which one can derive the HJB-equation of DP.
E. DP for the risk-sensitive problem

- For \( t \in [0, T] \) and \( Z_t = z, \xi_t = \xi, V_t = \upsilon \) put

\[
J_t(z, \xi, \upsilon; H) := \upsilon^\mu G_t(z, \xi; H)
\]

with

\[
G_t(z, \xi; H) = \tilde{E} \left\{ \exp \left[ -\mu \int_t^T \eta_s(Z_s, \xi_s, H_s) ds \right] \middle| Z_t = z, \xi_t = \xi \right\}
\]

- In view of the HJB equation put

\[
w(t, z, \xi) := \sup_H \log G_t(z, \xi; H)
\]

so that

\[
\sup_H J_0(z, \xi, \upsilon; H) = \upsilon^\mu e^{w(0, z, \xi)}
\]

\[ \rightarrow \text{ Notice that } w(T, z, \xi) = 0. \]
• Putting $Y_t = [Z_t, \xi_t]'$ and $B_t(Y_t) \equiv B_t(Z_t)$, one can synthesize the problem on $(\Omega, \mathcal{F}, \mathcal{F}_t, \tilde{P})$ into

$$
\begin{cases}
  dY_t = \left[ \Gamma_t(Y_t) + \mu H_t B_t(Y_t) \Sigma_t(Y_t) \right] dt + \Sigma_t(Y_t) d\tilde{w}_t \\
  \sup_H G_t(y, H); \quad y = [z, \xi]' 
\end{cases}
$$

• Recalling

$$
\eta_t(z, \xi, H) := \frac{1 - \mu}{2} H^2 B_t^2(z) - H \left( A_t(z, \xi) + \frac{1}{2} B_t^2(z) \right)
$$

the HJB equation for $w(t, z, \xi) := \sup_H \log G_t(z, \xi; H)$ is (formally)

$$
\begin{cases}
  w_t + \frac{1}{2} tr \left[ \Sigma_t(y) \Sigma'_t(y) w_{yy} \right] + \frac{1}{2} w'_y \Sigma_t(y) \Sigma'_t(y) w_y \\
  + \sup_H \left\{ \left( \Gamma_t(y) + \mu H B_t(y) \Sigma_t(y) \right)' w_y \right. \\
  + \left. \mu H \left[ A_t(y) + \frac{1}{2} B_t^2(y) \right] - \frac{\mu(1-\mu)}{2} H^2 B_t^2(y) \right\} = 0 \\
  w(T, y) = 0
\end{cases}
$$

$$
\Rightarrow \quad H_t^*(y) = \frac{1}{(1-\mu) B_t(y)} \Sigma'_t(y) w_y(t, y) \\
+ \frac{1}{(1-\mu) B_t^2(y)} \left[ A_t(y) + \frac{1}{2} B_t^2(y) \right]
$$
Replacing $H_t^*(y)$ for $H$ in the HJB equation one obtains the nonlinear 2nd order PDE

\[
\begin{cases}
  w_t + \frac{1}{2} tr \left[ \Sigma_t(y) \Sigma_t'(y) w_{yy} \right] \\
  \quad + \frac{1}{2(1-\mu)} w'_y \Sigma_t(y) \Sigma_t'(y) w_y + \Phi'(t,y) w_y + \Psi(t,y) = 0 \\
  w(T,y) = 0
\end{cases}
\]

where

\[
\begin{align*}
\Phi(t,y) &= \Gamma_t(y) + \frac{\mu}{(1-\mu) B_t(y)} \Sigma_t(y) \left[ A_t(y) + \frac{1}{2} B_t^2(y) \right] \\
\Psi(t,y) &= \frac{\mu}{2(1-\mu) B_t^2(y)} \left[ A_t(y) + \frac{1}{2} B_t^2(y) \right]^2
\end{align*}
\]

→ Putting $v(t,t) := \exp \left[ \frac{1}{1-\mu} w(t,y) \right]$, for $v(t,y)$ we obtain the following linear 2nd order PDE

\[
\begin{cases}
  v_t + \frac{1}{2} tr [\Sigma_t(y) \Sigma_t'(y) v_{yy}] + \Phi'(t,y) v_y + \frac{\Psi(t,y)}{1-\mu} v = 0 \\
  v(T,y) = 1
\end{cases}
\]
\( v(t, y) \) and thus also \( w(t, y) \) are unique viscosity solutions and \( v(t, y) \) can be computed as the expectation

\[
v(t, y) = E \left\{ \exp \left[ \frac{1}{1 - \mu} \int_t^T \Psi(s, Y_s) ds \right] | Y_t = y \right\}
\]

where \( Y_t \) now satisfies

\[
\begin{aligned}
\frac{dY_t}{dt} &= \Phi(t, Y_t) dt + \sum_t(Y_t) dw_t \\
Y_t &= y
\end{aligned}
\]

with \( w_t \) a Wiener process (uniqueness in distribution suffices).

\( \text{The latter representation may be used to compute } v(t, y) \text{ by simulation.} \)

Under sufficient regularity on the coefficients, \( v(t, y) \) and thus also \( w(t, y) \) are classical solutions implying that the optimal strategy (based on \( w_y \)) exists as a function of \( y = (z, \xi) \).
III. A pure jump model
(Di Masi, Callegaro, R. 05)

- Market model \( (B_t \equiv 1) \)

\[
dS^j_t = S^j_{t-} \left[ \sum_{i=1}^{M} (e^{a_{ij}} - 1) dN^i_t \right], \quad j = 1, \ldots, N
\]

\[\Rightarrow S^j_t = S^j_0 \exp \left[ \sum_{i=1}^{M} a_{ij} N^i_t \right]\]

- Self-financing portfolio (no consumption)

\[
\frac{dV_t}{V_{t-}} = \sum_{j=1}^{N} h^j_t \frac{dS^j_t}{S^j_{t-}} = \sum_{j=1}^{N} h^j_t \sum_{i=1}^{M} (e^{a_{ij}} - 1) dN^i_t
\]

implies

\[
V_T = V_0 \prod_{i=1}^{M} \exp \left[ \int_{0}^{T} \log \left( 1 + \sum_{j=1}^{N} h^j_t (e^{a_{ij}} - 1) \right) dN^i_t \right]
\]

which in turn leads to

\[
\begin{align*}
\log V_T &= \log V_0 + \sum_{i=1}^{M} \int_{0}^{T} \log \left( 1 + \sum_{j=1}^{N} h^j_t (e^{a_{ij}} - 1) \right) dN^i_t \\
V^\mu_T &= V^\mu_0 \prod_{i=1}^{M} \exp \left[ \int_{0}^{T} \log \left( 1 + \sum_{j=1}^{N} h^j_t (e^{a_{ij}} - 1) \right)^\mu dN^i_t \right]
\end{align*}
\]
Intensity of a pure jump process

• Let $N_t$ be a Poisson process adapted to $\mathcal{F}_t$

  $\rightarrow$ the $(P, \mathcal{F}_t)$–intensity $\lambda_t$ of $N_t$ is then characterized by

  $$E \left\{ \int_0^t C_s dN_s \right\} = E \left\{ \int_0^t C_s \lambda_s ds \right\}, \forall t \geq 0$$

  and $\forall$ nonnegative $\mathcal{F}_t$–predictable $C_t$.

  $\rightarrow$ Furthermore,

  $M_t := N_t - \int_0^t \lambda_s ds$

  is a $(P, \mathcal{F}_t)$–local martingale.
Let $\mathcal{F}_t^N \subseteq \mathcal{G}_t \subseteq \mathcal{F}_t$ and let $\lambda_t$ be the $(P, \mathcal{F}_t)$–intensity of $N_t$

$\rightarrow$ The process

$$\hat{\lambda}_t := E \{ \lambda_t \mid \mathcal{G}_t \}$$

is then the $(P, \mathcal{G}_t)$–intensity of $N_t$, i.e.

$$E \left\{ \int_0^t C_s dN_s \right\} = E \left\{ \int_0^t C_s \hat{\lambda}_s ds \right\} , \ \forall t \geq 0$$

and $\forall$ nonnegative $\mathcal{G}_t$–predictable $C_t$.

$\rightarrow$ Furthermore,

$$\hat{M}_t := N_t - \int_0^t \hat{\lambda}_s ds$$

is a $(P, \mathcal{G}_t)$–local martingale

$\rightarrow$ corresponds to the innovations process.
A. log-utility

A.1. Complete information (intensity known)

\[ \sup_h E \{ \log V_T \} = \log V_0 \]

\[ + \sup_h \sum_{i=1}^{M} E \left\{ \int_0^T \log \left( 1 + \sum_{j=1}^{N} h_t^j (e^{a_{ij}} - 1) \right) \lambda_t^i dt \right\} \]

→ It suffices to maximize for each \( t \)

\[ \sum_{i=1}^{M} \log \left( 1 + \sum_{j=1}^{N} h_t^j (e^{a_{ij}} - 1) \right) \lambda_t^i \]

→ Necessary condition

\[ \sum_{i=1}^{M} \frac{\lambda_t^i (e^{a_{ij}} - 1)}{1 + \sum_{j=1}^{N} h_t^j (e^{a_{ij}} - 1)} = 0; \quad j = 1, \ldots, N \]

→ For \( M = 2 \) one obtains a linear system of eqns. For \( M > 2 \) the system becomes non-linear.
A.2. Incomplete information \((\lambda_t \text{ is not observed, the observable filtration is } \mathcal{F}_t^S = \mathcal{F}_t^N)\)

- In this case \(h_t \in \mathcal{F}_t^S\)
- We can then write

\[
\sup_h E \{\log V_T\} = \log V_0 \\
+ \sup_h \sum_{i=1}^M E \left\{ \int_0^T \log \left(1 + \sum_{j=1}^N h_j^i (e^{a_{ij}} - 1)\right) \right\} \\
\cdot E \left\{ \lambda_t^i | \mathcal{F}_t^S \right\} dt
\]

\rightarrow \text{ A CE-principle holds in the sense that the unobserved } \lambda_t \text{ has simply to be replaced by } \hat{\lambda}_t = E \left\{ \lambda_t^i | \mathcal{F}_t^S \right\}. \]
B. Power utility

\[ E \left\{ V_T^\mu \right\} \]

\[ = V_0^\mu E \left\{ \prod_{i=1}^M \exp \left[ \int_0^T \log \left( 1 + \sum_{j=1}^N h_t^j(e^{a_{ij}} - 1) \right)^\mu dN_t^i \right] \right\} \]

\[ = V_0^\mu E \left\{ \prod_{i=1}^M \exp \left[ \int_0^T \log \left( 1 + \sum_{j=1}^N h_t^j(e^{a_{ij}} - 1) \right)^\mu dN_t^i \right. \right. \]

\[ + \int_0^T \left[ 1 - \left( 1 + \sum_{j=1}^N h_t^j(e^{a_{ij}} - 1) \right)^\mu \right] \lambda^i_t dt \]

\[ - \int_0^T \left[ 1 - \left( 1 + \sum_{j=1}^N h_t^j(e^{a_{ij}} - 1) \right)^\mu \right] \lambda^i_t dt \} \]

\[ = V_0^\mu \tilde{E} \left\{ \prod_{i=1}^M \exp \left[ \int_0^T \left[ (1 + \sum_{j=1}^N h_t^j(e^{a_{ij}} - 1))^\mu - 1 \right] \lambda^i_t dt \right] \right\} \]

\[ \rightarrow \tilde{E} \] corresponds to a measure \( \tilde{P} \sim P \) under which the intensity of \( N_t^i \) is

\[ \lambda^i_t \left( 1 + \sum_{j=1}^N h_t^j(e^{a_{ij}} - 1) \right)^\mu \]

\[ \rightarrow A \text{ CE-principle holds automatically: just do } \text{ the same exercise as above with } \tilde{\lambda}^i_t \text{ instead of } \lambda^i_t, (h_t^j \in \mathcal{F}^N). \]