Risk Averse Dynamic Programming

Andrés Ferragut

Universidad ORT Uruguay

UTE 2020

Contenido

Definiciones

Se considera un espacio de probabilidad $(\Omega, \mathcal{F}, P)$ en el cual están definidas variables aleatorias $X : \Omega \to \mathbb{R}$. Sea $\mathcal{X}$ el espacio vectorial $L^p = \{X : E[|X|^p] < \infty\}$ con $p \geqslant 1$.

Definición: Función de Riesgo

Una función de riesgo es un funcional $\rho : \mathcal{X} \to \mathbb{R}$ que asigna a cada variable aleatoria un valor de riesgo. Asumimos que X es un costo por lo que valores altos de X implican mayor riesgo.

Definición: Riesgo Coherente

Una función de riesgo se denomina coherente si verifica:

A1 **Convexidad:** $\rho(\lambda X + (1-\lambda)Y)) \leqslant \lambda\rho(X) + (1-\lambda)\rho(Y), \forall X, Y \in \mathcal{X}, \lambda \in [0,1]$.

A2 **Monotonía:** Para todo $X, Y \in \mathcal{X}$, $X \leqslant Y \Rightarrow \rho(X) \leqslant \rho(Y)$.

A3 **Equivariancia traslacional:** Para todo $X \in \mathcal{X}$, $c \in \mathbb{R}$, $\rho(X + c) = \rho(X) + c$.

A4 **Homogeneidad positiva:** Para todo $X \in \mathcal{X}$, $\lambda \geqslant 0$, $\rho(\lambda X) = \lambda\rho(X)$.

Ejemplos

Ejemplo (Valor esperado)

Para cada X

$$\rho(X) = E[X]$$

es una función de riesgo coherente.

No-ejemplo (Value at Risk)

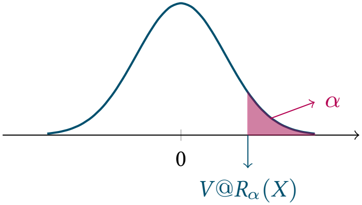
Para cada X

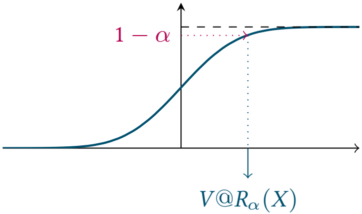
$$\rho(X) = V@R_\alpha[X] = \inf\{x : P(X > x) \leqslant \alpha\}$$

no *es una función de riesgo coherente. Verifica (A2-A4) pero no la convexidad.*

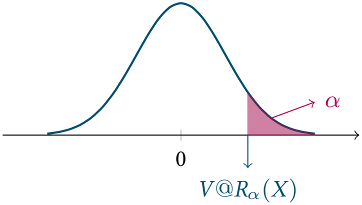
Value at Risk: interpretación

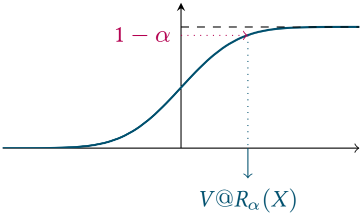
Caso continuo:



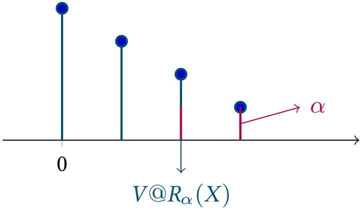


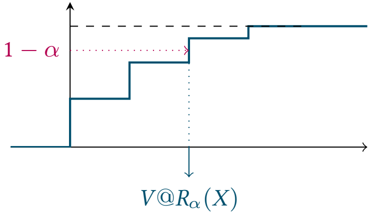
Caso continuo:





Caso discreto:



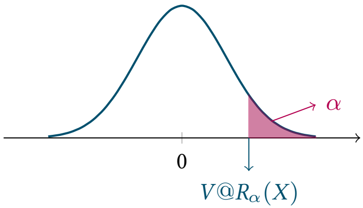


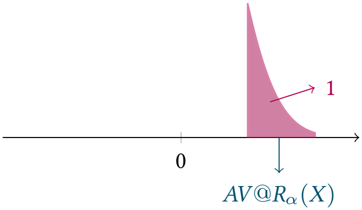
Average Value-at-Risk

Idea: Definir un valor medio de costo para el $\alpha\%$ peor de los casos.

Caso continuo:

$$AV@R_\alpha(X) = E[X \mid X > V@R_\alpha(X)] = \frac{1}{\alpha}\int_{V@R_\alpha(X)}^{\infty} xf(x)dx.$$



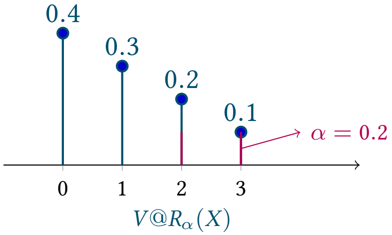


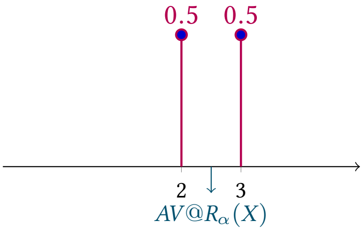
Average Value-at-Risk

Caso discreto: el problema está en que el cuantil α no es exacto, es decir en general:

$$P(X > V@R_\alpha(X)) < \alpha, \qquad P(X \geqslant V@R_\alpha(X)) > \alpha.$$

Idea: Considerar parte del átomo que se encuentra en el $V@R_\alpha(X)$ en el cálculo.





Average Value-at-Risk

Definición: Average value at risk

Para cada $X \in \mathcal{X}$ se define:

$$\rho(X) = AV@R_\alpha(X) = \inf_x \left\{ x + \frac{1}{\alpha} E[(X - x)^+] \right\}.$$

Average Value-at-Risk

Definición: Average value at risk

Para cada $X \in \mathcal{X}$ se define:

$$\rho(X) = AV@R_\alpha(X) = \inf_x \left\{ x + \frac{1}{\alpha} E[(X-x)^+] \right\}.$$

Propiedades:

- Es una medida de riesgo coherente (verifica A1-A4).
- Es conservadora respecto al $V@R_\alpha$, es decir: $AV@R_\alpha(X) \geqslant V@R_\alpha(X)$.

En el caso continuo la definición coincide con la que ya vimos. Sea:

$$g(x) = x + \frac{1}{\alpha} E[(X - x)^+]$$

Observando que $\frac{d}{dx}(X - x)^+ = -\mathbf{1}_{\{X>x\}}$ (ctp), se tiene que:

$$g'(x) = 1 + \frac{1}{\alpha} E[-\mathbf{1}_{\{X>x\}}] = 1 - \frac{1}{\alpha} P(X > x) = 0 \Leftrightarrow x = x^* =: V@R_\alpha(X).$$

De donde:

$$\begin{aligned} AV@R_\alpha(X) &= x^* + \frac{1}{\alpha} E[(X - x^*)^+] \\ &= \frac{1}{\alpha} E[(x^* + (X - x^*)^+)\mathbf{1}_{\{X \geqslant x^*\}}] \\ &= \frac{1}{\alpha} E[X \mathbf{1}_{\{X > x^*\}}] = E[X \mid X > V@R_\alpha(X)] \end{aligned}$$

Average Value-at-Risk

Caso discreto

Sea X una v.a. discreta con recorrido $x_1 < \ldots < x_n$ y probabilidades $P(X = x_i) = p_i$. Entonces:

$$g(x) = x + \frac{1}{\alpha}E[(X-x)^+] = x + \frac{1}{\alpha}\sum_{x_i > x} p_i(x_i - x).$$

La anterior es una función lineal a tramos de x. Nuevamente se tiene:

$$g'(x) = 1 - \frac{1}{\alpha}\sum_{x_i > x} p_i = 1 - \frac{1}{\alpha}\sum_{i} p_i \mathbf{1}_{\{x_i > x\}}.$$

La función anterior cambia de signo en $x^* : P(X > x^*) < \alpha$ y $P(X \geqslant x^*) \geqslant \alpha$, es decir, en $x^* = V@R_\alpha(x)$.

Evaluando en x^* obtenemos:

$$
\begin{aligned}
AV@R_\alpha(X) &= x^* + \frac{1}{\alpha} \sum_{x_i > x^*} p_i (x_i - x^*) \\
&= \frac{1}{\alpha} \left[\alpha x^* + \sum_{x_i > x^*} p_i (x_i - x^*) \right] \\
&= \frac{1}{\alpha} \left[\left(\alpha - \sum_{x_i > x^*} p_i \right) x^* + \sum_{x_i > x^*} p_i x_i \right] \\
&= \left(1 - \sum_{x_i > x^*} \frac{p_i}{\alpha} \right) x^* + \sum_{x_i > x^*} \frac{p_i}{\alpha} x_i
\end{aligned}
$$

Average Value-at-Risk

Caso discreto

Es decir, para calcular el $AV@R_\alpha(X)$ discreta debemos:

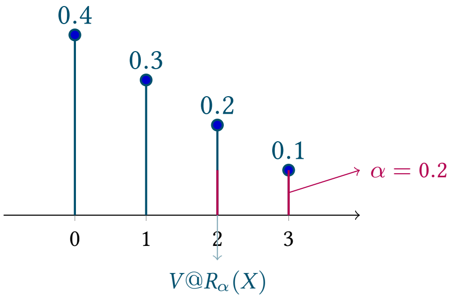
- Ordenar el recorrido $x_1 < x_2 < \ldots < x_n$.
- Hallar $x^* = V@R_\alpha(X)$ tal que $P(X > x^*) < \alpha$ y $P(X \geqslant x^*) \geqslant \alpha$.
- Reasignar las probabilidades del recorrido como:

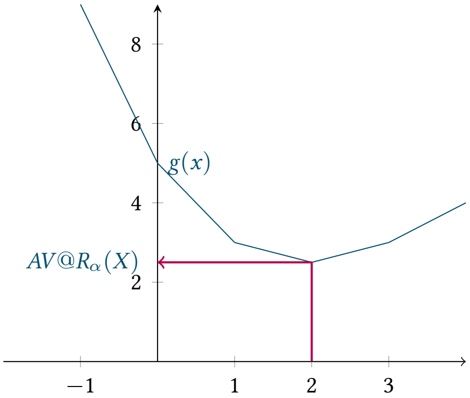
$$x_i < x^* \to 0, \quad x^* \to \left(1 - \sum_{x_i > x_*} \frac{p_i}{\alpha}\right), \quad x_i > x^* \to \frac{p_i}{\alpha}.$$

y en este caso $AV@R_\alpha(X) = E_\alpha[X]$ siendo E_α la esperanza con la nueva medida de probabilidad.

Average Value-at-Risk

Caso discreto: ejemplo





Contenido

Stochastic dynamic programming para el costo medio

Problema del costo medio

$$\min_{\pi} E\left[\sum_{k=0}^{N-1} g_k(x_k, u_k, w_k) + g_N(x_N)\right],$$

$$\begin{aligned} \text{sujeto a: } x_{k+1} &= f_k(x_k, u_k, w_k) \\ u_k &= \pi_k(x_k, w_k). \end{aligned}$$

con la hipótesis de que w_k son v.a. *iid*.

La esperanza anterior es en todos los w_k futuros.

Algoritmo de Bellman para el costo medio

Asumiendo un esquema ruido $\to$ decisión:

$$
\begin{aligned}
V_N(x) &= g_N(x),\\
\hat{V}_k(x, w) &= \min_{u_k}\{g_k(x, u_k, w) + V_{k+1}(f_k(x, u_k, w))\},\\
V_k(x) &= E\left[\hat{V}_k(x, w_k)\right]
\end{aligned}
$$

Política óptima:

$$
\pi_k(x, w) = \arg\min_{u_k}\{g_k(x, u_k, w) + V_{k+1}(f_k(x, u_k, w))\}.
$$

Problema de aversión al riesgo

Si ahora tenemos una medida de riesgo ρ uno podría plantear:

Problema global de aversión al riesgo

$$\min_{\pi} \rho \left[\sum_{k=0}^{N-1} g_k(x_k, u_k, w_k) + g_N(x_N) \right],$$

$$\begin{aligned} \text{sujeto a: } x_{k+1} &= f_k(x_k, u_k, w_k) \\ u_k &= \pi_k(x_k, w_k). \end{aligned}$$

con w_k *iid*.

Problema: No es posible descomponer aditivamente ρ, por lo que la estrategia de Bellman no funciona.

Problema de aversión al riesgo anidado

Como sustituto del problema anterior, uno puede plantear:

Problema anidado de aversión al riesgo

$$\min_{u_0} \rho_0 \left[g_0(x_0, u_0, w_0) + \rho_1 \left[\min_{u_1} g_1(x_1, u_1, w_1) + \rho_2[\cdots] \right] \right]$$

con las restricciones adicionales:

$$x_{k+1} = f_k(x_k, u_k, w_k)$$

y con w_k *iid*.

Nota: Las medidas de riesgo ρ_k toman en cuenta la información disponible hasta k (riesgo condicional al estado x_k y ruido w_k).

Stochastic dynamic programming con aversión al riesgo

El problema anterior admite una descomposición tipo Bellman:

Algoritmo de Bellman con aversión al riesgo

Asumiendo un esquema ruido $\to$ decisión:

$$
\begin{aligned}
V_N(x) &= g_N(x), \\
\hat{V}_k(x, w) &= \min_{u_k} \{ g_k(x, u_k, w) + V_{k+1}(f_k(x, u_k, w)) \}, \\
V_k(x) &= \rho_k \left[\hat{V}_k(x, w_k) \right]
\end{aligned}
$$

Política óptima:

$$\pi_k(x, w) = \arg\min_{u_k} \{ g_k(x, u_k, w) + V_{k+1}(f_k(x, u_k, w)) \}.$$

Contenido

Operador de Bellman para el costo medio

- Para cada k y una función $V : \mathcal{X} \times \mathcal{W} \to \mathbb{R}$ se define:

$$[\hat{T}_k(V)](x, w) := \min_u \{g_k(x, u, w) + V(f(x, u, w))\}.$$

- Con esta notación la iteración de Bellman resulta:

$$\begin{aligned} V_N(x) &= g_N(x), \\ V_k(x) &= T_k(V_{k+1}(x)) := E\left[\hat{T}(V_{k+1})(x, w_k)\right] \end{aligned}$$

- Este operador es monótono y preserva la convexidad.

SDDP para el costo medio

- Supongamos que tenemos una cota inferior $V_{k+1}^{(l+1)} \leqslant V_{k+1}$ y un estado $x_k^{(l)}$. Resolvemos:

$$\begin{aligned} \hat{\beta}_k^{(l+1)}(w) = \min_{x,u} & \left\{ g_k(x,u,w) + V_{k+1}^{(l+1)}(f_k(x,u,w)) \right\}, \\ s.t. \quad & x = x_k^{(l)} \quad [\hat{\lambda}_k^{(l+1)}(w)]. \end{aligned}$$

- Por lo tanto:

$$\hat{\beta}_k^{(l+1)}(w) = \hat{T}_k(V_{k+1}^{(l+1)})(x_k^{(l)}, w),$$
$$\hat{\lambda}_k^{(l+1)}(w) \in \partial_x \hat{T}_k(V_{k+1}^{(l+1)})(x_k^{(l)}, w).$$

- Para cada perturbación entonces:

$$\hat{\beta}_k^{(l+1)}(w) + (\hat{\lambda}_k^{(l+1)})^T(x - x_k^{(l)}) \leqslant \hat{T}_k(V_{k+1}^{(l+1)})(x, w) \leqslant \hat{V}_k(x, w).$$

SDDP para el costo medio

- Promediando sobre los escenarios, tenemos una cota inferior para la función de valor:
$$\beta_k^{(l+1)} + (\lambda_k^{(l+1)})^T(x - x_k^{(l)}) \leqslant E(\hat{V}_k(x, w)) = V_k(x).$$
- En cada paso definimos:

$$\begin{aligned}
\beta_k^{(l+1)} &= E[\hat{\beta}_k^{(l+1)}(w)] = T_k(V_{k+1}^{(l+1)})(x), \\
\lambda_k^{(l+1)} &= E[\hat{\lambda}_k^{(l+1)}(w)] \in \partial_x T_k(V_{k+1}^{(l+1)})(x).
\end{aligned}$$

lo que define un nuevo corte para la aproimación de V_k.

SDDP con aversión al riesgo

Supongamos que disponemos de una medida de aversión al riesgo $\rho(X)$ tal que:

$$\rho(X) = E_P[X]$$

siendo P una medida de probabilidad adecuada.

- Ejemplos:
  - $\rho(X) = E[X]$.
  - $\rho(X) = AV@R_\alpha[X] = E_\alpha[X]$ con E_α como vimos antes.
  - $\rho(X) = \lambda E[X] + (1-\lambda)AV@R_\alpha[X]$, $\lambda \in [0,1]$ permite "matizar" la aversión al riesgo.
- ¡Entonces toda la deducción anterior vale cambiando E por E_P!

Algoritmo SDDP con aversión al riesgo

Al comienzo de cada paso: disponemos de una aproximación $V_k^{(l)}$ de V_k tal que:

- $V_k^{(l)} \leqslant V_k$.
- $V_N^{(l)} = g_N$, el costo terminal.
- $V_k^{(l)}$ es convexa (mejor aún, lineal a tramos...)

- Seleccionamos un escenario al azar $w_0, \ldots, w_{N-1}$.

- Construimos la trayectoria $x_k^{(l)}$ siguiendo la dinámica:

$$
\begin{aligned}
u_k^{(l)} &= \arg\min_u \left\{ g_k(x_k^{(l)}, u, w_k) + V_{k+1}^{(l)}(f(x_k^{(l)}, u, w_k) \right\}, \\
x_{k+1}^{(l)} &= f_k(x_k^{(l)}, u_k^{(l)}, w_k).
\end{aligned}
$$

- Es decir, la trayectoria calculada como si las funciones de cost-to-go se fueran sus aproximaciones $V_k^{(l)}$.

Algoritmo SDDP

Backward iteration:

- En cada k queremos mejorar la estimación de $V_k^{(l)}$.
- Resolvemos, en cada k, para todo posible w:

$$\hat{\beta}_k^{(l+1)}(w) = \min_{x,u}\{g_k(x,u,w) + V_{k+1}^{(l+1)}(f_k(x,u,w))\},$$
$$s.t. \quad x = x_k^{(l)} \quad [\hat{\lambda}_k^{(l+1)}(w)].$$

- Con la lista de valores $\{\hat{\beta}_k^{(l+1)}(w_i), i = 1, \dots, M\}$, construimos la medida de probabilidad P que permite calcular $\rho\left[\hat{\beta}_k^{(l+1)}(w)\right]$ como una esperanza.

Ejemplo: si ρ es el $AV@R_\alpha$, ordenamos los valores de menor a mayor, calculamos el cuantil α y hacemos la transformación de probabilidades ya vista.

Algoritmo SDDP

Backward iteration:

- Calculamos el subgradiente y óptimo promedio:

$$\beta_k^{(l+1)} = E_P[\hat{\beta}_k^{(l+1)}(w)], \qquad \lambda_k^{(l+1)} = E_P[\hat{\lambda}_k^{(l+1)}(w)].$$

  - Es decir, se promedia utilizando las probabilidades transforamadas tanto en β como en λ.
- Agregamos el corte a la estimación:

$$V_k^{(l+1)}(x) = \max\{V_k^{(l)}(x), \beta_k^{(l+1)} + (\lambda_k^{(l+1)})^T(x - x_k^{(l)})\}$$

- Retrocedemos de k a $k-1$ y al llegar a 0 se completa la pasada.

Contenido

Ejemplo: el problema del vendedor de periódicos

Consideremos el siguiente problema:

- Un vendedor de periódicos debe decidir cuántos reservar. El precio de reserva (el día antes) es p.
- El día de venta recibe una demanda de periódicos $W \sim U[0, \theta]$ (para fijar ideas).
- Si la demanda supera la reserva, debe comprar periódicos extra a precio $q > p$.

Preguntas:

- ¿Cuánto debemos reservar si queremos minimizar el costo medio?
- ¿Cuánto debemos reservar si queremos minimizar el $AV@R_\alpha$ del costo?

Tomemos:

- $x_0 = 0$ el stock inicial.
- u_0 el control (reserva).
- $g_0(x_0, u_0) = pu_0$, costo del primer paso.
- $x_1 = x_0 + u_0$ la dinámica. x_1 es el stock al día siguiente.
- $g_1(x_1, u_1, w_1) = qu_1$.
- Con la restricción adicional de que $x_2 = x_1 + u_1 - w_1 \geqslant 0$ (debo cumplir toda la demanda en la segunda compra), además $u_0 \geqslant 0$, $u_1 \geqslant 0$.

Ejemplo: el problema del vendedor de periódicos

Newsvendor problem, average cost formulation

Hallar:

$$\min_u \left\{pu + E[q(W-u)^+]\right\}$$

con $W \sim U[0,\theta]$ en este caso.

Newsvendor problem, risk averse formulation

Hallar:

$$\min_u \left\{pu + \rho_\alpha[q(W-u)^+]\right\}$$

con $W \sim U[0,\theta]$ en este caso.

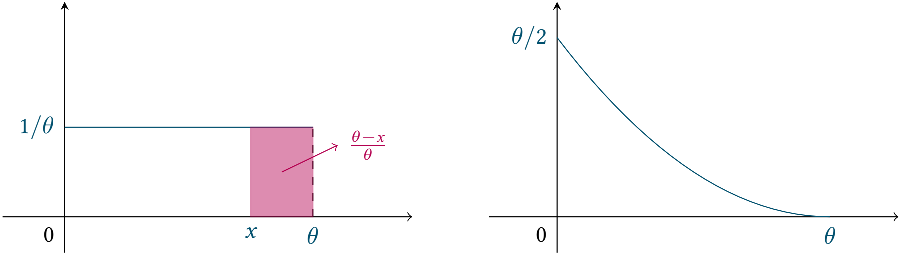
Observación: $V(x) = V(u) = E[q(W-x)^+]$ o $V_\alpha(x) = V_\alpha(u) = \rho_\alpha[q(W-x)^+]$ juegan el rol de función de costo/riesgo futuro.

Cálculo del costo futuro

Haciendo cuentas con la $U[0,\theta]$ se llega a:

- Costo medio futuro:

$$E\left[(W-x)^+\right] = \frac{(\theta-x)^2}{2\theta}$$

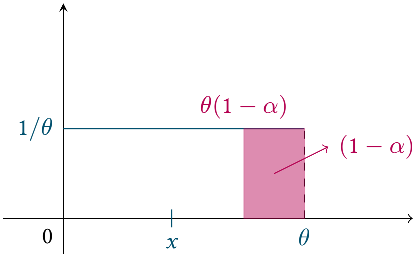


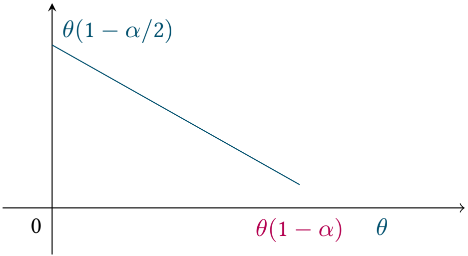
Cálculo del costo futuro

Nuevamente haciendo cuentas con la $U[0,\theta]$ se llega a:

- Riesgo medio futuro:

$$\rho_\alpha\left[(W-x)^+\right] = \begin{cases} \theta\left(1-\frac{\alpha}{2}\right) - x, & \text{si } x < \theta(1-\alpha) \\ \frac{(\theta-x)^2}{2\alpha\theta}, & \text{si } x \geqslant \theta(1-\alpha) \end{cases}$$



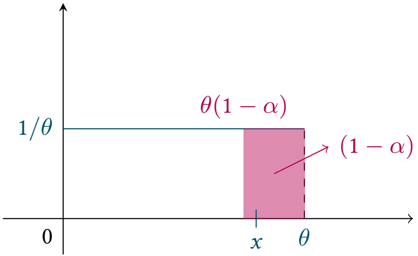


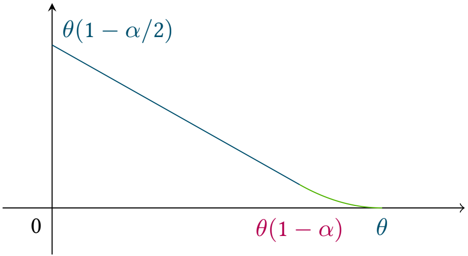
Cálculo del costo futuro

Nuevamente haciendo cuentas con la $U[0,\theta]$ se llega a:

- Riesgo medio futuro:

$$\rho_\alpha\left[(W-x)^+\right] = \begin{cases} \theta\left(1-\frac{\alpha}{2}\right)-x, & \text{si } x<\theta(1-\alpha) \\ \frac{(\theta-x)^2}{2\alpha\theta}, & \text{si } x \geqslant \theta(1-\alpha) \end{cases}$$





Solución del problema de costo medio

Debemos hallar:

$$\min_u \{pu + qE[(W-u)^+]\}$$

con $p < q$, que corresponde a:

$$\min_u \left\{ pu + q\frac{(\theta - u)^2}{2\theta} \right\}$$

Derivando e igualando a 0, obtenemos:

$$u^* = \theta\left(1 - \frac{p}{q}\right).$$

Solución del problema de aversión al riesgo

Debemos hallar:

$$\min_u \{pu + q\rho_\alpha[(W-u)^+]\}$$

con $p < q$. Derivando $q\rho_\alpha$ se tiene que la parte lineal tiene pendiente $p - q < 0$ por lo que el mínimo se da después de $\theta(1-\alpha)$. Derivando la segunda parte se llega a la condición:

$$p - q\frac{\theta - u}{\alpha\theta} = 0,$$

por lo que la compra óptima es:

$$u^* = \theta\left(1 - \alpha\frac{p}{q}\right).$$

Resolución mediante SDDP

En cuaderno de Julia adjunto

Implementación en Julia

- La biblioteca SDDP.jl permite correr SDDP con aversión al riesgo.
- Requiere ruidos discretos, y puede considerar el caso de incertidumbre markoviana.
- Incluye $AV@R_\alpha$ y una combinación convexa de esperanza y $AV@R$.
- Notar que si $\alpha \to 0$ corresponde a optimizar el peor caso, lo que también está incluido.
- No conlleva una penalidad de performance (misma cantidad de LPs, agrega un sort para construir la medida de probabilidad del $AV@R$).
- En preparación: cuaderno de Julia con un ejemplo hidrotérmico.

Muchas gracias!

Contacto: ferragut@ort.edu.uy
Código y slides disponibles en: https://github.com/Grupo-MATE/risk_averse_sddp