Optimisation under probability constraints: an approach via quantiles

Felisa J. Vázquez-Abad

Joint work with Pierre Carpentier (ENSTA, France) and Guy Cohen (ENPC, France) Honours project Andrea Macrae (Melbourne)

> Department of Mathematics and Statistics University of Melbourne, Australia

Conference in Honour of Reuven's Birthday, 14-18 July 2008

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

▲ 同 ▶ → 三 ▶

## Probability Constraints

• Model qualitative risk: fatal failure or death (if we eat a bad cheese, it does not matter how much we ate beyond the fatal dose)

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

- Model qualitative risk: fatal failure or death (if we eat a bad cheese, it does not matter how much we ate beyond the fatal dose)
- Chemical processes control pressure, temperature, etc

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

- Model qualitative risk: fatal failure or death (if we eat a bad cheese, it does not matter how much we ate beyond the fatal dose)
- Chemical processes control pressure, temperature, etc
- Financial investments to control risk of ruin, shortage of funds, etc

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

- Model qualitative risk: fatal failure or death (if we eat a bad cheese, it does not matter how much we ate beyond the fatal dose)
- Chemical processes control pressure, temperature, etc
- Financial investments to control risk of ruin, shortage of funds, etc
- Telecommunication networks to control loss of information, loss probability, error rates, etc

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

- Model qualitative risk: fatal failure or death (if we eat a bad cheese, it does not matter how much we ate beyond the fatal dose)
- Chemical processes control pressure, temperature, etc
- Financial investments to control risk of ruin, shortage of funds, etc
- Telecommunication networks to control loss of information, loss probability, error rates, etc
- Service industry to control measures of client satisfaction

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

- 4 同 2 4 日 2 4 日 2

### **Probability** Constraints

• Problem:  $g(\cdot, u) \colon \mathbb{R} \to \mathbb{R}$ ,  $\xi$  a continuous rv min J(u) s.t.  $\mathbb{P}(g(\xi, u) \le \alpha) \ge p$ .

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

- 4 同 2 4 日 2 4 日 2

## **Probability** Constraints

• Problem:  $g(\cdot, u) \colon \mathbb{R} \to \mathbb{R}$ ,  $\xi$  a continuous rv min J(u) s.t.  $\mathbb{P}(g(\xi, u) \le \alpha) \ge p$ .

• Let 
$$\zeta(u) = g(\xi, u)$$
, then

$$\mathbb{P}\{g(\xi, u) \leq lpha\} \geq p \quad \Rightarrow \quad \mathbb{P}\{\zeta(u) \leq lpha\} \geq p,$$

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

- 4 同 2 4 日 2 4 日 2

## **Probability** Constraints

• Problem:  $g(\cdot, u) \colon \mathbb{R} \to \mathbb{R}$ ,  $\xi$  a continuous rv min J(u) s.t.  $\mathbb{P}(g(\xi, u) \le \alpha) \ge p$ .

• Let 
$$\zeta(u) = g(\xi, u)$$
, then

$$\mathbb{P}\{ m{g}(\xi, m{u}) \leq lpha \} \geq m{p} \quad \Rightarrow \quad \mathbb{P}\{\zeta(m{u}) \leq lpha \} \geq m{p},$$

• Constraint 
$$B(u) = p - F_{\zeta(u)}(\alpha) = p - F[g^{-1}(\alpha, u)]$$

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

< 白⇒ < (

-

## **Probability** Constraints

• Problem:  $g(\cdot, u) \colon \mathbb{R} \to \mathbb{R}$ ,  $\xi$  a continuous rv min J(u) s.t.  $\mathbb{P}(g(\xi, u) \le \alpha) \ge p$ .

• Let 
$$\zeta(u) = g(\xi, u)$$
, then

$$\mathbb{P}\{m{g}(\xi,m{u})\leqlpha\}\geqm{p} \quad \Rightarrow \quad \mathbb{P}\{\zeta(m{u})\leqlpha\}\geqm{p},$$

• Constraint 
$$B(u) = p - F_{\zeta(u)}(\alpha) = p - F[g^{-1}(\alpha, u)]$$

### **Distribution Formulation**

$$\min_{u} J(u)$$
  
subject to:  $B(u) \leq 0$ 

Felisa J. Vázquez-Abad Probability Constraints

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

・ 同・ ・ ヨ・

## Example

- We borrow one unit \$\$ at interest rate I, pay at end of period
- Decision: fraction  $u_1$  to invest at fixed rate b < l
- Decision: fraction u<sub>2</sub> to invest at risky rate ξ, E[ξ] > I
- Consumption is  $1 u_1 u_2$ , "utility" or satisfaction from consumption is  $U(\cdot)$  concave non decreasing.

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

<ロト < 同ト < 三ト

## Example

- We borrow one unit \$\$ at interest rate I, pay at end of period
- Decision: fraction  $u_1$  to invest at fixed rate b < l
- Decision: fraction  $u_2$  to invest at risky rate  $\xi, \mathbb{E}[\xi] > I$
- Consumption is  $1 u_1 u_2$ , "utility" or satisfaction from consumption is  $U(\cdot)$  concave non decreasing.

s.t.  
$$\begin{split} \max_{u_1,u_2} \mathbb{E} \big( U(1-u_1-u_2) + (1+b)u_1 + (1+\xi)u_2 \big) \\ & u \geq 0 \,, \quad u_1+u_2 \leq 1 \,, \\ \mathbb{P} \big( (1+b)u_1 + (1+\xi)u_2 \geq 1 + l \big) \geq p \,. \end{split}$$

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

イロト イポト イヨト イヨト

## Example: the challenges

Challenges with the problem include:

 Non-linear optimisation problem of the form min<sub>u</sub> J(u), s.t. B(u) ≤ 0.

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

(4 同 ) (1 日 ) (1 日 )

## Example: the challenges

- Non-linear optimisation problem of the form  $\min_u J(u)$ , s.t.  $B(u) \le 0$ .
- Black or Grey box models: input (u, ξ) and output g(ξ, u), J(u), J'(u), but distribution of ξ unknown: how to use statistical estimation for simulation-based optimisation?

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

(4 同 ) (4 回 ) (4 回 )

## Example: the challenges

- Non-linear optimisation problem of the form  $\min_u J(u)$ , s.t.  $B(u) \le 0$ .
- Black or Grey box models: input (u, ξ) and output
   g(ξ, u), J(u), J'(u), but distribution of ξ unknown: how to use
   statistical estimation for simulation-based optimisation?
- Gradient-based algorithms (stochastic) require

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

## Example: the challenges

- Non-linear optimisation problem of the form  $\min_u J(u)$ , s.t.  $B(u) \le 0$ .
- Black or Grey box models: input (u, ξ) and output
   g(ξ, u), J(u), J'(u), but distribution of ξ unknown: how to use
   statistical estimation for simulation-based optimisation?
- Gradient-based algorithms (stochastic) require
  - convexity of  $B(u) = p \mathbb{P}(g(\xi, u) \le \alpha)$

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

- 4 同 2 4 日 2 4 日 2

## Example: the challenges

- Non-linear optimisation problem of the form  $\min_u J(u)$ , s.t.  $B(u) \le 0$ .
- Black or Grey box models: input (u, ξ) and output g(ξ, u), J(u), J'(u), but distribution of ξ unknown: how to use statistical estimation for simulation-based optimisation?
- Gradient-based algorithms (stochastic) require
  - convexity of  $B(u) = p \mathbb{P}(g(\xi, u) \le \alpha)$
  - estimation of gradient of a probability

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

## Example: the challenges

- Non-linear optimisation problem of the form  $\min_u J(u)$ , s.t.  $B(u) \le 0$ .
- Black or Grey box models: input (u, ξ) and output
   g(ξ, u), J(u), J'(u), but distribution of ξ unknown: how to use
   statistical estimation for simulation-based optimisation?
- Gradient-based algorithms (stochastic) require
  - convexity of  $B(u) = p \mathbb{P}(g(\xi, u) \le \alpha)$  !!!!
  - estimation of gradient of a probability

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

イロト イポト イヨト イヨト

## Example: the challenges

- Non-linear optimisation problem of the form  $\min_u J(u)$ , s.t.  $B(u) \leq 0$ .
- Black or Grey box models: input (u, ξ) and output g(ξ, u), J(u), J'(u), but distribution of ξ unknown: how to use statistical estimation for simulation-based optimisation?
- Gradient-based algorithms (stochastic) require
  - convexity of  $B(u) = p \mathbb{P}(g(\xi, u) \le \alpha)$  !!!!
  - estimation of gradient of a probability (discontinuities, lack of model for distribution)

Problem Formulation

Research Question Contributions Example Concluding Remarks Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

### Example: Optimal Cost



Figure: Optimal cost as a function of "confidence" level p.

Non-convexity and non saturated but active constraints.

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

< A > <

## **Constrained** Optimisation

#### Theorem

For a convex problem (strictly convex J(u) and B(u)) the optimal  $(u^*, \lambda^*)$  is a saddle point and solves:

$$\min_{u\in\mathbb{R}^d}\max_{\lambda\geq 0}L(u,\lambda)=\max_{\lambda\geq 0}\min_{u\in\mathbb{R}^d}L(u,\lambda)$$

The Arrow Hurwicz Algorithm is:

$$u_{n+1} = u_n - \epsilon_n \left( \nabla_u J(u_n) + \lambda_n^T \nabla_u B(u_n) \right)$$
$$\lambda_{n+1} = \max(0, \lambda_n + \epsilon_n B(u_{n+1}))$$

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

< A > <

## **Constrained** Optimisation

#### Theorem

For a convex problem (strictly convex J(u) and B(u)) the optimal  $(u^*, \lambda^*)$  is a saddle point and solves:

$$\min_{u\in\mathbb{R}^d}\max_{\lambda\geq 0}L(u,\lambda)=\max_{\lambda\geq 0}\min_{u\in\mathbb{R}^d}L(u,\lambda)$$

The Arrow Hurwicz Algorithm is:

$$u_{n+1} = u_n - \epsilon_n \left( \nabla_u J(u_n) + \lambda_n^T \nabla_u B(u_n) \right) \text{grad min}$$
  
$$\lambda_{n+1} = \max(0, \lambda_n + \epsilon_n B(u_{n+1}))$$

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

< A > <

## **Constrained** Optimisation

#### Theorem

For a convex problem (strictly convex J(u) and B(u)) the optimal  $(u^*, \lambda^*)$  is a saddle point and solves:

$$\min_{u\in\mathbb{R}^d}\max_{\lambda\geq 0}L(u,\lambda)=\max_{\lambda\geq 0}\min_{u\in\mathbb{R}^d}L(u,\lambda)$$

The Arrow Hurwicz Algorithm is:

$$u_{n+1} = u_n - \epsilon_n \left( \nabla_u J(u_n) + \lambda_n^T \nabla_u B(u_n) \right) \text{grad min}$$
  
$$\lambda_{n+1} = \max(0, \lambda_n + \epsilon_n B(u_{n+1})) \text{ grad max}$$

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

The Arrow-Hurwicz Algorithm: Convergence

#### Theorem

Let  $x_{n+1} = x_n + \epsilon V(x_n)$  and let  $x_{\epsilon}(t) = x_n$ ,  $t \in [n\epsilon, (n+1)\epsilon)$ . If V is a Lipschitz continuous and bounded function, then as  $\epsilon \to 0$ ,  $x_{\epsilon}(\cdot)$  converges (in the sup norm) to the solution of the ODE:

$$\frac{dx(t)}{dt} = V[x(t)]$$

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

The Arrow-Hurwicz Algorithm: Convergence

#### Theorem

Let  $x_{n+1} = x_n + \epsilon V(x_n)$  and let  $x_{\epsilon}(t) = x_n$ ,  $t \in [n\epsilon, (n+1)\epsilon)$ . If V is a Lipschitz continuous and bounded function, then as  $\epsilon \to 0$ ,  $x_{\epsilon}(\cdot)$  converges (in the sup norm) to the solution of the ODE:

$$\frac{dx(t)}{dt} = V[x(t)]$$

• Convexity is not required for this property to hold.

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

The Arrow-Hurwicz Algorithm: Convergence

#### Theorem

Let  $x_{n+1} = x_n + \epsilon V(x_n)$  and let  $x_{\epsilon}(t) = x_n$ ,  $t \in [n\epsilon, (n+1)\epsilon)$ . If V is a Lipschitz continuous and bounded function, then as  $\epsilon \to 0$ ,  $x_{\epsilon}(\cdot)$  converges (in the sup norm) to the solution of the ODE:

$$\frac{dx(t)}{dt} = V[x(t)]$$

- Convexity is not required for this property to hold.
- Local convergence around stable points: study only the behaviour of active constraints: λ > 0 (continuity).

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

The Arrow-Hurwicz Algorithm: Convergence

#### Theorem

Let  $x_{n+1} = x_n + \epsilon V(x_n)$  and let  $x_{\epsilon}(t) = x_n$ ,  $t \in [n\epsilon, (n+1)\epsilon)$ . If V is a Lipschitz continuous and bounded function, then as  $\epsilon \to 0$ ,  $x_{\epsilon}(\cdot)$  converges (in the sup norm) to the solution of the ODE:

$$\frac{dx(t)}{dt} = V[x(t)]$$

- Convexity is not required for this property to hold.
- Local convergence around stable points: study only the behaviour of active constraints:  $\lambda > 0$  (continuity).
- Allows to characterise behaviour around stationary points.

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

## The A-H Algorithm: Convergence and Optimality.

- The vector field of A-H has stable points that are saddlepoints of the Lagrangian (convex problems).
- Linearise around a stable point  $x^*$  using Taylor expansion

$$V(x) = V(x^*) + \mathbb{A}(x - x^*) + (O)(||x - x^*||^2),$$
  
 $x(t) - x^* \approx e^{\mathbb{A}t}, \quad \mathbb{A} = \nabla V(x^*)^T$ 

A Hurwitz: ℜ(eigenv(A)) < 0 then x<sup>\*</sup> attractor (limit).

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

< ロ > < 同 > < 回 > <

## The A-H Algorithm: Convergence and Optimality.

- The vector field of A-H has stable points that are saddlepoints of the Lagrangian (convex problems).
- Linearise around a stable point  $x^*$  using Taylor expansion

$$V(x) = V(x^*) + \mathbb{A}(x - x^*) + (O)(||x - x^*||^2),$$
  

$$x(t) - x^* \approx e^{\mathbb{A}t}, \quad \mathbb{A} = \nabla V(x^*)^T$$

• A Hurwitz:  $\Re(eigenv(\mathbb{A})) < 0$  then  $x^*$  attractor (limit).

#### Result

Convex problem,  $x^* = (u^*, \lambda^*)$  optimal, constraint qualification  $\nabla B(x^*)$  l.i. vector. Then A is Hurwitz, implying that the optimal solution and multiplier are attractors of the ODE.

Probability Constraints Example Constrained Optimisation The Arrow-Hurwicz Algorithm: Lagrange Duality

## The A-H Algorithm: Convergence and Optimality.

- The vector field of A-H has stable points that are saddlepoints of the Lagrangian (convex problems).
- Linearise around a stable point  $x^*$  using Taylor expansion

$$V(x) = V(x^*) + \mathbb{A}(x - x^*) + (O)(||x - x^*||^2),$$
  

$$x(t) - x^* \approx e^{\mathbb{A}t}, \quad \mathbb{A} = \nabla V(x^*)^T$$

• A Hurwitz:  $\Re(eigenv(\mathbb{A})) < 0$  then  $x^*$  attractor (limit).

### Result

Convex problem,  $x^* = (u^*, \lambda^*)$  optimal, constraint qualification  $\nabla B(x^*)$  l.i. vector. Then A is Hurwitz, implying that the optimal solution and multiplier are attractors of the ODE.

Non-convex problems: V(x) for insight into algorithm behaviour.

Vector Fields Arrow-Hurwitz

(4月) (4日) (4日)

### Vector Fields: examples

# Example $\min \frac{1}{2}u^2$ , s.t. $\mathbb{P}(\xi - u \le \alpha) \ge p$ $(u^0 = 0)$

$$B(u) = p - F(u + \alpha).$$

Case 1: uniform distribution

$$F(\xi) = 0.5 + 0.5(\xi - 0.5) \mathbf{1}_{\{-0.5 < \xi \le 0.5\}}$$

Case 2: "beta"-like distribution

$$F(\xi) = 0.5 + 0.5(2\xi)^3 \mathbf{1}_{\{-0.5 < \xi \le 0.5\}}$$

Vector Fields Arrow-Hurwitz

### Vector Fields: examples



Felisa J. Vázquez-Abad

Vector Fields Arrow-Hurwitz

### Vector Fields: examples



Felisa J. Vázquez-Abad

Vector Fields Arrow-Hurwitz

### Vector Fields: examples



- Depicted: Case 2.
- Unconstrained optimum at  $u^0 = 0$ .
- Feasible region is  $F(u+1) \ge 0.7$

- ▲ 🖓 🕨 - ▲ 🖻

picture aside

Felisa J. Vázquez-Abad

Vector Fields Arrow-Hurwitz

### Vector Fields: examples



- Depicted: Case 2.
- Unconstrained optimum at  $u^0 = 0$ .
- Feasible region is  $F(u+1) \ge 0.7$
- Solution is  $u^* \approx 1.36 \neq u^0$ .

picture aside

Felisa J. Vázquez-Abad

Vector Fields Arrow-Hurwitz

### Vector Fields: examples



- Depicted: Case 2.
- Unconstrained optimum at  $u^0 = 0$ .
- Feasible region is  $F(u+1) \ge 0.7$
- Solution is  $u^* \approx 1.36 \neq u^0$ .

< 17 ▶

• Probability constraint is active at optimum.

picture aside

Vector Fields Arrow-Hurwitz

Image: A image: A

### Vector Fields: examples



Figure: Left: Case 1: convex. Right: Case 2: non convex.

Vector Fields Arrow-Hurwitz

### Vector Fields: examples



Figure: Left: Case 1: convex. Right: Case 2: non convex.

Contribution

Conjecture: lack of convexity is the problem.

Overview of Contributions Pathology from bounded suppor Quantile formulation

<ロト < 同ト < 三ト

B> B

Contribution

**Overview of Contributions** Pathology from bounded support Quantile formulation

э

Conjecture: lack of convexity is the problem. ... or is it?

**Overview of Contributions** Pathology from bounded support Quantile formulation

- ▲ 🖓 🕨 - ▲ 🖻

## Contribution

Conjecture: lack of convexity is the problem. ... or is it?



Figure: Case 1: convex.

**Overview of Contributions** Pathology from bounded support Quantile formulation

## Contribution



Figure: Convex distribution: zoom out

**Overview of Contributions** Pathology from bounded support Quantile formulation

э



• Conjecture: lack of convexity is the problem.

**Overview of Contributions** Pathology from bounded support Quantile formulation

・ 同・ ・ ヨ・



- Conjecture: lack of convexity is the problem.
- Identification of problem: distributions with bounded support (potential numerical problem for any distribution)

**Overview of Contributions** Pathology from bounded support Quantile formulation

< A >



- Conjecture: lack of convexity is the problem.
- Identification of problem: distributions with bounded support (potential numerical problem for any distribution)
- Re statement of problem using a Quantile Formulation.

Overview of Contributions Pathology from bounded support Quantile formulation

## Pathology from bounded support

- For each u,  $g(\cdot, u)$  is monotone increasing,  $h(u, \xi) = g_u^{-1}(x)$ .
- $g(\cdot, u)$  is continuously differentiable in u.
- Bounded support  $F(\xi) = 0$ , for all  $\xi \leq \underline{\xi}$  and assume that  $\mathcal{U} = \{u \colon h(u, a) \leq \underline{\xi}\} \neq \emptyset$

### Theorem

Assume a unique optimal solution  $(u^*, \lambda^*)$  to the constrained problem

min 
$$J(u)$$
 s.t.  $B(u) = p - \mathbb{P}(g(\xi, u) \le \alpha) \le 0$ .

and that the unconstrained minimum  $u^0 = \arg \min_u J(u) \in \mathcal{U}$ . Then the A-H algorithm diverges when initialising "close" to  $u^0$ ; specifically,  $u_n \to u^0 \neq u^*$  and  $\lambda_n \to +\infty$ .

Overview of Contributions Pathology from bounded support Quantile formulation

< ロ > < 同 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < 回 > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ > < □ >

## Pathology from bounded support

### Proof.

• A-H algorithm has a vector field:

$$u' = -J'(u) + \lambda f(g^{-1}(\alpha, u))(g^{-1}(\alpha, u))$$
$$\lambda' = (p - F(g^{-1}(\alpha, u))\mathbf{1}_{\{\lambda \ge 0\}}$$

When initialising inside U, F(u) = f(u) = 0 so the algorithm behaves:

$$u' = -J'(u) \qquad \Rightarrow \qquad u \to u^0$$
  
 $\lambda' = p \qquad \Rightarrow \qquad \lambda \to +\infty$ 

Overview of Contributions Pathology from bounded support Quantile formulation

< fi ▶ <

## Quantile Formulation

#### Remark

Common methods to deal with no convexity can be used (penalties, augmented Lagrangian, A-H "beta" method for convexification, etc), but they will also suffer from the pathology of bounded support.

Overview of Contributions Pathology from bounded support Quantile formulation

## Quantile Formulation

### Remark

Common methods to deal with no convexity can be used (penalties, augmented Lagrangian, A-H "beta" method for convexification, etc), but they will also suffer from the pathology of bounded support.

- Fact: if distribution function is convex (concave) then its inverse the quantile function is concave (convex)
- Conjecture: use one or another to deal with regions of non convexity
- But our results show that Quantile formulation always works! (under convexity of  $g(\xi, \cdot)$ ).

Overview of Contributions Pathology from bounded support Quantile formulation

イロト イポト イヨト イヨト

## **Quantile Formulation**

#### Lemma

Suppose that for every u,  $g(\cdot, u)$  is monotone increasing. Then

 $\mathbb{P}(g(\xi, u) \leq \alpha) \geq p \Leftrightarrow g(Q(p), u) \leq \alpha.$ 

Overview of Contributions Pathology from bounded support Quantile formulation

▲□ ► < □ ► </p>

## Quantile Formulation

#### Lemma

Suppose that for every u,  $g(\cdot, u)$  is monotone increasing. Then

 $\mathbb{P}(g(\xi, u) \leq \alpha) \geq p \Leftrightarrow g(Q(p), u) \leq \alpha.$ 

- Suppose u is such that  $g(Q(p), u) \leq \alpha$
- By monotonicity. for all  $\xi \leq Q(p), g(\xi, u) \leq \alpha$ )
- Thus  $\mathbb{P}(g(\xi, u) \leq \alpha) = \mathbb{P}(\xi \leq Q(p)) \geq p$

Overview of Contributions Pathology from bounded support Quantile formulation

< (T) >

## Quantile Formulation

#### Lemma

Suppose that for every u,  $g(\cdot, u)$  is monotone increasing. Then

 $\mathbb{P}(g(\xi, u) \leq \alpha) \geq p \Leftrightarrow g(Q(p), u) \leq \alpha.$ 

- Suppose u is such that  $g(Q(p), u) \leq \alpha$
- By monotonicity. for all  $\xi \leq Q(p), g(\xi, u) \leq \alpha$ )
- Thus  $\mathbb{P}(g(\xi, u) \leq \alpha) = \mathbb{P}(\xi \leq Q(p)) \geq p$

### Quantile Formulation

min J(u) s.t.  $g(Q(p), u) \leq \alpha$ 

Overview of Contributions Pathology from bounded support Quantile formulation

< A > <

## **Quantile Formulation**

min 
$$J(u)$$
 s.t.  $g(Q(p), u) \le \alpha$   
 $L(u, \lambda) = J(u) + \lambda(g(Q(p), u) - \alpha).$ 

#### Theorem

If  $J(\cdot)$  and  $g(x, \cdot)$  are convex for every x, where x is a continuous random variable, then independently of the distribution function, AH has a unique attractor at the optimum.

Overview of Contributions Pathology from bounded support Quantile formulation

< A > <

## Quantile Formulation

min 
$$J(u)$$
 s.t.  $g(Q(p), u) \le \alpha$   
 $L(u, \lambda) = J(u) + \lambda(g(Q(p), u) - \alpha).$ 

#### Theorem

If  $J(\cdot)$  and  $g(x, \cdot)$  are convex for every x, where x is a continuous random variable, then independently of the distribution function, AH has a unique attractor at the optimum.

$$\dot{u_t} = -\nabla J(u_t) - \nabla_u g(Q(p), u_t)$$
  
 $\dot{\lambda_t} = (g(Q(p), u) - \alpha) \mathbb{1}_{\{\lambda_t \ge 0\}}$ 

Overview of Contributions Pathology from bounded support Quantile formulation

Image: A mathematic states and a mathematic states

## Quantile Formulation

min 
$$J(u)$$
 s.t.  $g(Q(p), u) \le \alpha$   
 $L(u, \lambda) = J(u) + \lambda(g(Q(p), u) - \alpha).$ 

#### Theorem

If  $J(\cdot)$  and  $g(x, \cdot)$  are convex for every x, where x is a continuous random variable, then independently of the distribution function, AH has a unique attractor at the optimum.

$$\dot{u_t} = -
abla J(u_t) - 
abla_u g(Q(p), u_t)$$
  
 $\dot{\lambda_t} = (g(Q(p), u) - lpha) \mathbf{1}_{\{\lambda_t \ge 0\}}$ 

If decreasing then use g(Q(1-p), u)

Overview of Contributions Pathology from bounded support Quantile formulation

### **Quantile Formulation**



### Figure: Quantile formulation

## Example

- We borrow one unit \$\$ at interest rate I, pay at end of period
- Decision: fraction  $u_1$  to invest at fixed rate b < l
- Decision: fraction u<sub>2</sub> to invest at risky rate ξ, E[ξ] > I
- Consumption is  $1 u_1 u_2$ , "utility" or satisfaction from consumption is  $U(\cdot)$  concave non decreasing.

## Example

- We borrow one unit \$\$ at interest rate I, pay at end of period
- Decision: fraction  $u_1$  to invest at fixed rate b < l
- Decision: fraction  $u_2$  to invest at risky rate  $\xi, \mathbb{E}[\xi] > I$
- Consumption is 1 − u<sub>1</sub> − u<sub>2</sub>, "utility" or satisfaction from consumption is U(·) concave non decreasing.

s.t.  
$$\begin{split} \max_{u_1,u_2} \mathbb{E} \big( U(1-u_1-u_2) + (1+b)u_1 + (1+\xi)u_2 \big) \\ & u \geq 0 \,, \quad u_1+u_2 \leq 1 \,, \\ \mathbb{P} \big( (1+b)u_1 + (1+\xi)u_2 \geq 1 + l \big) \geq p \,. \end{split}$$

< A >

### **Example:** Quantile Formulation

Here the constraint function is decreasing:  $g(u_1, u - 2, \xi) = -(1 + b)u_1 - (1 + \xi)u_2, a = l + 1$ , so we use:  $B(u_1, u_2) = (l + 1) - (1 + b)u_1 - (1 + Q(1 - p))u_2$ .

・ 同 ト ・ ヨ ト ・ ヨ ト

### **Example:** Quantile Formulation

Here the constraint function is decreasing:  $g(u_1, u - 2, \xi) = -(1 + b)u_1 - (1 + \xi)u_2, a = l + 1, \text{ so we use:}$   $B(u_1, u_2) = (l + 1) - (1 + b)u_1 - (1 + Q(1 - p))u_2.$   $\dot{u_1} = -U'(1 - u_1 - u_2) - \lambda_1 b - \lambda_2$   $\dot{u_2} = -U'(1 - u_1 - u_2) - \lambda_1 Q(1 - p) - \lambda_2$   $\dot{\lambda_1} = (l + 1) - (1 + b)u_1 - (1 + \xi)u_2$   $\dot{\lambda_2} = u_1 + u_2 - 1.$ 

Fast convergence to optimal point, no problem for the algorithm. Note that now the multiplier gives sensitivity w.r.t. the level of constraint a(-l+1) rather than to p.

- 4 同 2 4 日 2 4 日 2 4

Concluding Remarks

## **On-going work**

- Quantile formulation promises better algorithmic behaviour.
- Can the formulation be extended to piecewise monotonic functions?
- How to use the approach for simulation: open question.
- How to generalise to several variables: open question.
- Current research with France: aerospace control, needs a dynamical system and g(·, u) depends on whole trajectory.

Concluding Remarks

## The End

- Thank you for your attention
- Questions?

3 x 3