THE SCENARIO APPROACH to
STOCHASTIC OPTIMIZATION

Marco C. Campi
University of Brescia
Italy
“What I like about experience is that it is such an honest thing. … You may have deceived yourself, but experience is not trying to deceive you.”

C.S. Lewis
thanks to:

Algo Care'

Simone Garatti

Giuseppe Calafiore

Maria Prandini

Bernardo Pagnoncelli

Federico Ramponi
optimization

- controller synthesis
- classification
- portfolio selection

}  optimization program
optimization

- controller synthesis
- classification
- portfolio selection

optimization program

uncertain environment

- exercise caution
uncertain optimization:

\[
\begin{align*}
\text{optimize } & J(\theta, \delta) \\
\end{align*}
\]
uncertain optimization:

\[
\text{optimize } J(\theta, \delta)_{\theta}
\]

not a valid mathematical formulation
uncertain optimization:

\[
\underset{\theta}{\text{optimize }} J(\theta, \delta)
\]

not a valid mathematical formulation

often, a description of uncertainty is not available, or it is only partially available
scenario-based knowledge:
scenario-based knowledge:

knowledge about uncertainty can be acquired through experience,

that is, we look at previous cases, or scenarios, of the same problem
example: portfolio optimization

[with B. Pagnoncelli & D. Reich]
example: portfolio optimization

[with B. Pagnoncelli & D. Reich]
example: portfolio optimization

1

d assets

[with B. Pagnoncelli & D. Reich]
example: portfolio optimization

1$

d$ assets

$p_k = \text{percentage of capital invested on asset } k$

[with B. Pagnoncelli & D. Reich]
example: portfolio optimization

$1$

d assets

$p_k = \text{percentage of capital invested on asset } k$

$\theta = [p_1 \cdots p_d]^T$
example: portfolio optimization

$J(\theta, \delta) = \sum_{k=1}^{d} p_k R_k$

$\theta = [p_1 \cdots p_d]^T$

$\delta = [R_1 \cdots R_d]^T$

$1$ assets

$p_k = \text{percentage of capital invested on asset } k$

[with B. Pagnoncelli & D. Reich]
example: portfolio optimization

[with B. Pagnoncelli & D. Reich]
example: portfolio optimization

record of past rate of returns:

\[ R_k(i) = \text{return of asset } k \text{ over period } i \]  (scenarios)

[with B. Pagnoncelli & D. Reich]
example: portfolio optimization

record of past rate of returns:

\[ R_k(i) = \text{return of asset } k \text{ over period } i \]  
(scenarios)

\[
J(\theta, \delta^{(i)}) = \sum_{k=1}^{d} p_k R_k(i), \quad i = 1 \ldots, N
\]

[with B. Pagnoncelli & D. Reich]
example: classification - defibrillation
example: classification - defibrillation

decide whether a defibrillator has to be applied

[with A. Caré]
example: classification - defibrillation

(scenarios)

[with A. Caré]
example: classification - defibrillation

(scenarios)

[with A. Caré]
“scenario” optimization (convex case)
min-max “scenario” optimization

[with G. Calafiore]
min-max “scenario” optimization

\[ J(\theta, \delta) \text{ convex in } \theta \]

[with G. Calafiore]
min-max “scenario” optimization

\[ J(\theta, \delta) \text{ convex in } \theta \]

\[ \delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)} \rightarrow J(\theta, \delta^{(1)}), J(\theta, \delta^{(2)}), \ldots, J(\theta, \delta^{(N)}) \]

[with G. Calafiore]
min-max “scenario” optimization

$J(\theta, \delta)$ convex in $\theta$

$\delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)} \rightarrow J(\theta, \delta^{(1)}), J(\theta, \delta^{(2)}), \ldots, J(\theta, \delta^{(N)})$

[with G. Calafiore]
**min-max “scenario” optimization**

\[ J(\theta, \delta) \quad \text{convex in } \theta \]

\[ \delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)} \quad \rightarrow \quad J(\theta, \delta^{(1)}), J(\theta, \delta^{(2)}), \ldots, J(\theta, \delta^{(N)}) \]

[with G. Calafiore]
min-max “scenario” optimization

\[ J(\theta, \delta) \text{ convex in } \theta \]

\[ \delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)} \rightarrow J(\theta, \delta^{(1)}), J(\theta, \delta^{(2)}), \ldots, J(\theta, \delta^{(N)}) \]

[with G. Calafiore]
min-max “scenario” optimization

\[ J(\theta, \delta) \text{ convex in } \theta \]

\[ \delta^{(1)}, \delta^{(2)}, \ldots, \delta^{(N)} \rightarrow J(\theta, \delta^{(1)}), J(\theta, \delta^{(2)}), \ldots, J(\theta, \delta^{(N)}) \]

[with G. Calafiore]

\[ \text{SP}_N = \text{Senario Program} \]
what are the guarantees for unseen situations?
what are the guarantees for unseen situations?

how guaranteed is $J^*$ for another $\delta$?
what are the guarantees for unseen situations?

how guaranteed is $J^*$ for another $\delta$?

from the “visible” to the “invisible”
about uncertainty
about uncertainty
about uncertainty
about uncertainty
about uncertainty
about uncertainty

\[ \text{prob} = \text{risk} \]
Theorem (with S. Garatti)

Fix \( \epsilon \in (0, 1) \) (risk parameter)
\[ \beta \in (0, 1) \] (confidence parameter)

If \( N = \frac{2}{\epsilon} \left( \ln \frac{1}{\beta} + d \right) \),
then,
with probability \( \geq 1 - \beta \),
risk \( \leq \epsilon \).
**Theorem** (with S. Garatti)

Fix $\epsilon \in (0, 1)$ (risk parameter)

If $N = \frac{2}{\epsilon} \left( \ln \frac{1}{\beta} + d \right)$,
then,

risk $\leq \epsilon$. 
Theorem (with S. Garatti)

Fix $\epsilon \in (0, 1)$ (risk parameter)
$\beta \in (0, 1)$ (confidence parameter)

If $N = \frac{2}{\epsilon} \left( \ln \frac{1}{\beta} + d \right)$,
then,
with probability $\geq 1 - \beta$,
risk $\leq \epsilon$. 
**Theorem** (with S. Garatti)

Fix $\epsilon \in (0, 1)$ (risk parameter)

If $N = \frac{2}{\epsilon}(7 \ln 10 + d)$,

then,

risk $\leq \epsilon$. 
comments

generalization \rightarrow \text{need for structure}

good news: the structure we need is only convexity
... more comments

\[ N = \frac{2}{\varepsilon} \left( \ln \frac{1}{\beta} + d \right) \]

- \( N \) depends on how complex the decision is via \( d \)
- \( N \) does not depend on how complex the “real world” is
more comments

\[ N = \frac{2}{\epsilon} \left( \ln \frac{1}{\beta} + d \right) \]

- \( N \) depends on how complex the decision is via \( d \)
- \( N \) does not depend on how complex the “real world” is

don’t try to reconstruct the real world to answer easy questions!
… more comments  (IPM = Interval Prediction Model)
... more comments  (IPM = Interval Prediction Model)
\( y = \theta_1 + \theta_2 u + \theta_3 u^2 + \theta_4 u^3 \)

\[
\min_{\theta_1, \theta_2, \theta_3, \theta_4} \left[ \max_i |y_i - [\theta_1 + \theta_2 u_i + \theta_3 u_i^2 + \theta_4 u_i^3]| \right]
\]

... more comments  (IPM = Interval Prediction Model)
... more comments  (IPM = Interval Prediction Model)

\[ y = \theta_1 + \theta_2 u + \theta_3 u^2 + \theta_4 u^3 \]

risk = prob. that next point is outside IPM
… more comments  (IPM = Interval Prediction Model)
... more comments

\[ N = \frac{2}{\epsilon} \left( \ln \frac{1}{\beta} + d \right) \]

- \( N \) is independent of \( Pr \) \hspace{1em} (distribution-free result)
$N = \frac{2}{\varepsilon} \left( \ln \frac{1}{\beta} + d \right)$

- $N$ is independent of $Pr$  (distribution-free result)

“What I like about experience is that it is such an honest thing. … You may have deceived yourself, but experience is not trying to deceive you.”

C.S. Lewis
a more general theoretical result
Theorem (with S. Garatti)

The “risk” has a beta-distribution (universal).
Theorem (with S. Garatti)

The “risk” has a beta-distribution (universal).
Theorem (with S. Garatti)

The “risk” has a beta-distribution (universal).
Theorem (with S. Garatti)

The “risk” has a beta-distribution (universal).

![Graph showing beta-distributions for different values of N.]
Theorem (with S. Garatti)

The “risk” has a beta-distribution (universal).
Theorem (with S. Garatti)

The "risk" has a beta-distribution (universal).
Theorem (with S. Garatti)

The “risk” has a beta-distribution (universal).
**Theorem** (with S. Garatti)

The “risk” has a beta-distribution (universal).
Theorem (with S. Garatti)

The “risk” has a beta-distribution (universal).

\[ E[\text{risk}] = \frac{d+1}{N+1} \]
Theorem (with S. Garatti)

The “risk” has a beta-distribution (universal).

\[ E[\text{risk}] = \frac{d+1}{N+1} = \text{total probability of error} \]
application: classification - defibrillation

# Instances = 170
# NoRosc = 155
# Rosc = 15
application: classification - defibrillation

# Instances = 170
# NoRosc = 155
# Rosc = 15

\[ d = \text{complexity of classifier} \]
application: classification - defibrillation

\# Instances = 170
\# NoRosc = 155
\# Rosc = 15

\[ d = \text{complexity of classifier} \]

10-fold cross-validation
application: classification - defibrillation

# Instances = 170
# NoRosc = 155
# Rosc = 15

d = complexity of classifier
10-fold cross-validation

<table>
<thead>
<tr>
<th>d</th>
<th>3</th>
<th>9</th>
<th>15</th>
<th>21</th>
<th>27</th>
</tr>
</thead>
<tbody>
<tr>
<td># of errors</td>
<td>4 (2.35%)</td>
<td>10 (5.88%)</td>
<td>18 (10.59%)</td>
<td>23 (13.53%)</td>
<td>27 (15.88%)</td>
</tr>
<tr>
<td># of unknowns</td>
<td>143</td>
<td>132</td>
<td>73</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>E[risk]</td>
<td>1.96%</td>
<td>5.88%</td>
<td>9.80%</td>
<td>13.73%</td>
<td>17.65%</td>
</tr>
</tbody>
</table>
generalizations and beyond
generalizations: risk-return tradeoff
generalizations: risk-return tradeoff
generalizations: risk-return tradeoff
generalizations: risk-return tradeoff
generalizations: risk-return tradeoff
generalizations: risk-return tradeoff
generalizations: risk-return tradeoff
Theorem (risk-return trade-off)

With probability $\geq 1 - \beta$, risk of $J_k^* \leq \epsilon_k$ where:

$$\epsilon_k = \frac{k}{N} + O\left(\frac{1}{\sqrt{N}}\right)$$
Theorem (risk-return trade-off)

With probability $\geq 1 - \beta$, risk of $J^*_k \leq \epsilon_k$ where:

$$\epsilon_k = \frac{k}{N} + O\left(\frac{1}{\sqrt{N}}\right)$$
performance - risk plot
performance - risk plot
generalizations

\[
\min_{\theta} \max_i J(\theta, \delta^{(i)})
\]
generalizations

\[
\min_{\theta} \max_i J(\theta, \delta^{(i)})
\]

\[
\min_{\theta, J} J
\]

subject to: \( J \geq J(\theta, \delta^{(i)}), \quad i = 1, \ldots, N \)
generalizations

\[ \min_{\theta} c^T \theta \]
subject to: \( \theta \in \Theta_i, \quad i = 1, \ldots, N \)
generalizations

\[ \min_{\theta} c^T \theta \]
subject to: \( \theta \in \Theta_i, \quad i = 1, \ldots, N \)

relevant to:  
- quantitative finance (minimum return)  
- control with constraints (MPC)
Many have given a contribution:

the theory is still in its infancy
the theory is still in its infancy

non-convex
the theory is still in its infancy
... concluding

the problem of extracting knowledge from observations is perhaps the most central issue of all science
… concluding

the problem of extracting knowledge from observations is perhaps the most central issue of all science

the scenario approach is one way, and a lot of work remains to be done
... concluding

the problem of extracting knowledge from observations is perhaps the most central issue of all science

the scenario approach is one way, and a lot of work remains to be done

certainly: it is a wonderful world to explore!
… concluding

the problem of extracting knowledge from observations is perhaps the most central issue of all science

the scenario approach is one way, and a lot of work remains to be done

certainly: it is a wonderful world to explore!

THANK YOU!
REFERENCES

M.C. Campi and S. Garatti.
*The Exact Feasibility of Randomized Solutions of Uncertain Convex Programs.*

M.C. Campi and S. Garatti.

B.K. Pagnoncelli, D. Reich and M.C. Campi
*Risk-Return Trade-off with the Scenario Approach: A Case study in Portfolio Selection.*

G. Calafiore and M.C. Campi.
*Uncertain Convex Programs: randomized Solutions and Confidence Levels.*

G. Calafiore and M.C. Campi.
*The Scenario Approach to Robust Control Design.*

M.C. Campi, G. Calafiore and S. Garatti.
*Interval Predictor Models: Identification and Reliability.*

M.C. Campi.
*Classification with guaranteed probability of error.*