DEFENCE R&D DÉFENSE

Sensitivity Analysis of Weighted-Sum Scoring Methods

Bohdan L. Kaluzny

Defence Research & Development Canada

Centre for Operational Research & Analysis

CORS Ottawa, 27 November 2009

Outline

1. Motivation
2. Previous work
3. Methodology
   - Geometric intuition
   - Sensitivity measures
4. Example
5. Conclusion

Motivation

- Buying a new car:
  - Narrowed down search to 4 options
    - Subaru, GM, Honda, BMW
  - Identified crucial evaluation criteria:
    - Price, fuel efficiency, passenger capacity,
    - safety rating, dealership proximity, warranty
  - How to proceed?

Criteria	Scoring function
C_1: Price ($K)	over $50K, 0 pts; under $20K, 10pts; otherwise ($50K - price)/3 pts.
C_2: Fuel efficiency (mpg)	over 50 mpg, 10 pts; otherwise (fuel efficiency)/5 pts.
C_3: Passenger capacity	over 6, 10 pts; otherwise (passenger capacity)/0.6 pts.
C_4: Safety (star rating)	(number of stars)$\times$2 pts.
C_5: Servicing proximity	over 10km away, 0 pts; otherwise (10km - distance) pts.
C_6: Warranty length (years)	over 10 years, 10 pts; otherwise (warranty length) pts.

Motivation

- Weighted-Sum scoring method

	Criteria	Subaru	GM	Honda	BMW
35	C_1 ($K)	32 (6 pts)	23 (9 pts)	26 (8 pts)	38 (4 pts)
15	C_2 (mpg)	30 (6 pts)	25 (5 pts)	40 (8 pts)	35 (7 pts)
10	C_3 (passengers)	5 (8.33 pts)	4 (6.67 pts)	5 (8.33 pts)	6 (10 pts)
21	C_4 (stars)	5 (10 pts)	4 (8 pts)	4.5 (9 pts)	5 (10 pts)
7	C_5 (kms)	10 (0 pts)	2 (8 pts)	4 (6 pts)	8 (2 pts)
12	C_6 (years)	3 (3 pts)	7 (7 pts)	3 (3 pts)	5 (5 pts)
		6.293	**7.647**	**7.503**	**6.410**

- Given n options and m criteria
  - v_{ij} = rating of option j relative to criterion i
  - W_i = weight allocated to criterion i
  - $S_j = \sum v_{ij} \cdot W_i$ = total score for option j

Motivation

- Weighted-Sum scoring method

		Criteria	Subaru	GM	Honda	BMW
31	~~35~~	C_1 ($K)	32 (6 pts)	23 (9 pts)	26 (8 pts)	38 (4 pts)
19	~~15~~	C_2 (mpg)	30 (6 pts)	25 (5 pts)	40 (8 pts)	35 (7 pts)
10	10	C_3 (passengers)	5 (8.33 pts)	4 (6.67 pts)	5 (8.33 pts)	6 (10 pts)
21	21	C_4 (stars)	5 (10 pts)	4 (8 pts)	4.5 (9 pts)	5 (10 pts)
7	7	C_5 (kms)	10 (0 pts)	2 (8 pts)	4 (6 pts)	8 (2 pts)
12	12	C_6 (years)	3 (3 pts)	7 (7 pts)	3 (3 pts)	5 (5 pts)
			~~6.293~~ 6.293	~~7.647~~ 7.487	~~7.503~~ 7.503	~~6.290~~ 6.410

- Given n options and m criteria
  - v_{ij} = rating of option j relative to criterion i
  - W_i = weight allocated to criterion i
  - $S_j = \sum v_{ij} \cdot W_i$ = total score for option j

Motivation

How typical are the chosen weights?

How sensitive is the final ranking to changes in these weights?

Has someone fine-tuned the weights?

Previous work

Several approaches…

Gahrlein and Fishburn (1983)

- Probability that two randomly selected weight vectors would yield the same rank ordering

Evans (1984)

- Geometrical 'maximum confidence sphere' around baseline weights yielding same rank ordering

Schneller and Sphicas (1984)

- Geometric analysis not new: **Starr's domain criterion (1962)**, work of **Isaacs (1963), (1965)**

Barron and Schmidt (1988)

- Entropy-based and least-squared methods to find nearest weights that change top-ranked option

Triantaphyllou and Sanchez (1997)

- Determining most critical criterion: smallest weight change results in altered ranking

Butler et al (1997)

- Monte-Carlo simulation for simultaneous variation of all weights

Morrice *et al.* (1999)

- One-at-a-time analysis of each weight holding the ratio of the other weights constant

Idea

We can define the sensitivity analysis problem geometrically

We can then use algorithms and results from high-dimensional computational geometry:

a sub-field known as **Polyhedral Computation**

Significant advances in the last 20 years in Polyhedral Computation: lrs, cdd, VINCI, CPLEX that allow us to analyze reasonably sized problems

Methodology

Geometric intuition:

Three options X, Y and Z.

Three criteria with weights w_1, w_2, w_3: $\sum_{i=1}^{3} w_i = 100$

$w_i \geq 0 \quad \text{for } i = 1, \ldots, 3$

Recall: $S_j = \sum_{i=1}^{m} v_{ij} w_i$ = total score for option j

When is option X top ranked?

→ Interested in **weight space** when $S_X > S_Y$ and $S_X > S_Z$

For what W's does the ranking < X Y Z > hold?

→ Interested in **weight space** when $S_X > S_Y > S_Z$

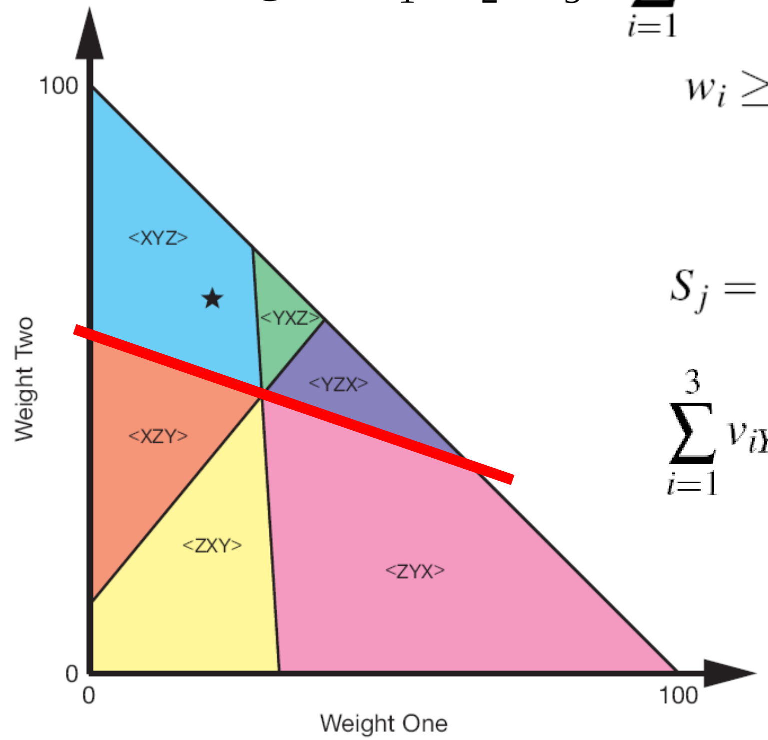
Methodology

Geometric intuition:

Three options X, Y and Z.

Three criteria with weights w_1, w_2, w_3: $\sum_{i=1}^{3} w_i = 100$

$$w_i \geq 0 \quad \text{for } i = 1, \ldots, 3$$



$$S_j = \sum_{i=1}^{m} v_{ij} w_i$$

$$\sum_{i=1}^{3} v_{iY} w_i - \sum_{i=1}^{3} v_{iZ} w_i \leq 0$$

Methodology

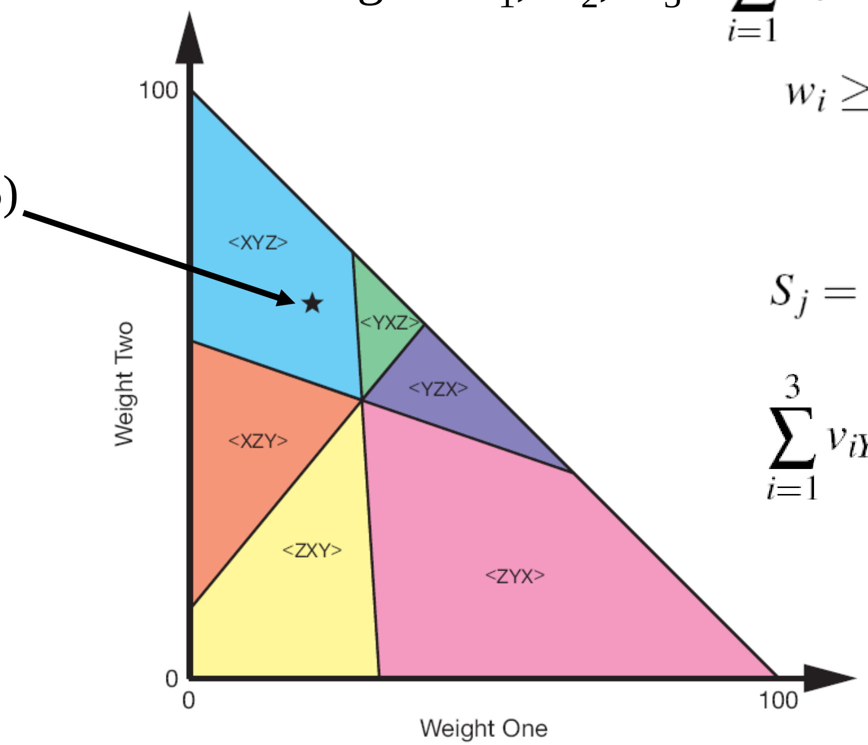
Geometric intuition:

Three options X, Y and Z.

Three criteria with weights w_1, w_2, w_3: $\sum_{i=1}^{3} w_i = 100$

$$w_i \geq 0 \quad \text{for } i = 1, \ldots, 3$$

W= (25, 60, 15)



$$S_j = \sum_{i=1}^{m} v_{ij} w_i$$

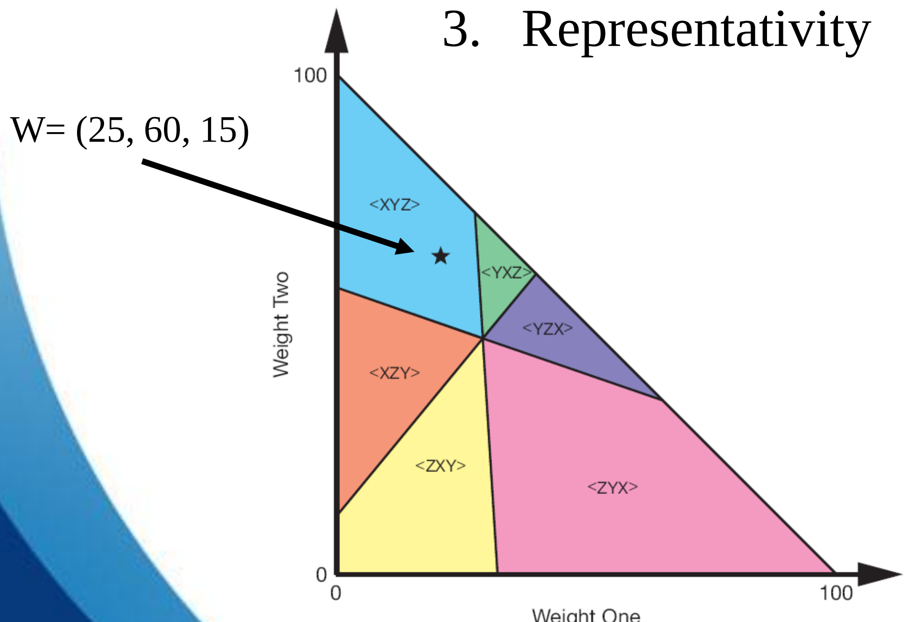
$$\sum_{i=1}^{3} v_{iY} w_i - \sum_{i=1}^{3} v_{iZ} w_i \leq 0$$

Methodology

Geometric intuition:

Sensitivity measures:

1. Distance-based
2. Volume-based
3. Representativity

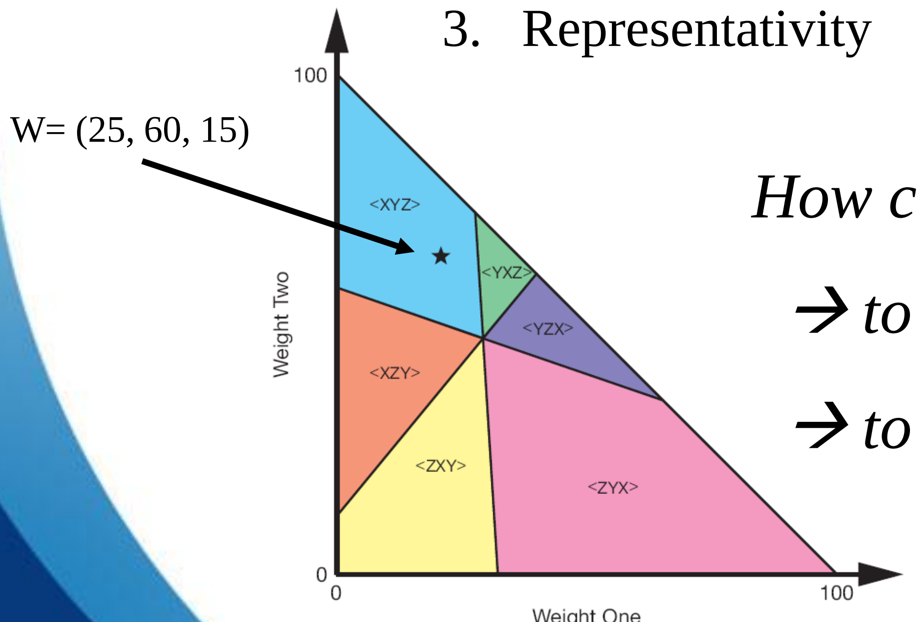


Methodology

Geometric intuition:

Sensitivity measures:

1. **Distance-based**
2. Volume-based
3. Representativity



How close is nearest boundary?

→ *to a weight being zero*

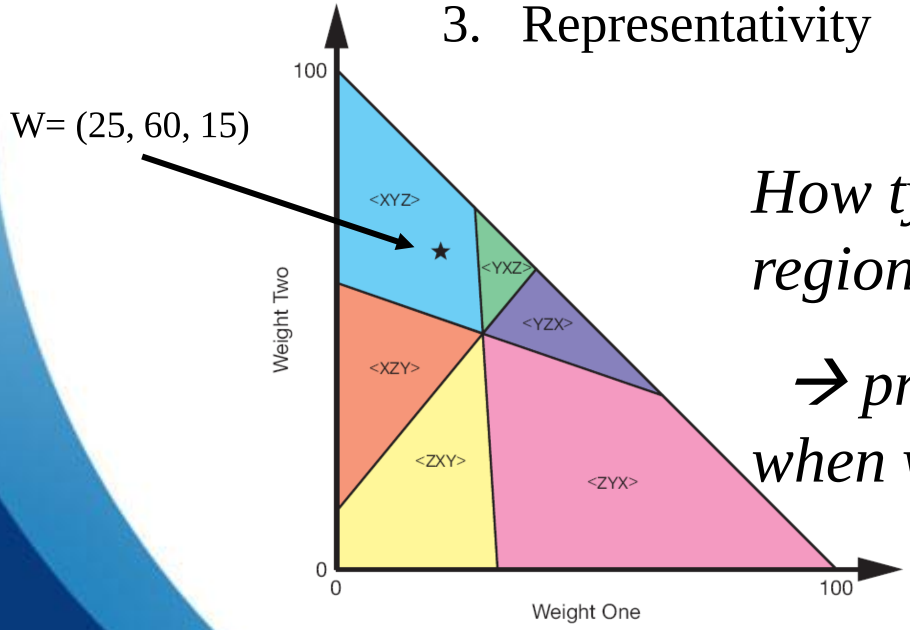
→ *to alternative ranking of options*

Methodology

Geometric intuition:

Sensitivity measures:

1. Distance-based
2. **Volume-based**
3. Representativity



How typical (large) is the ranking region?

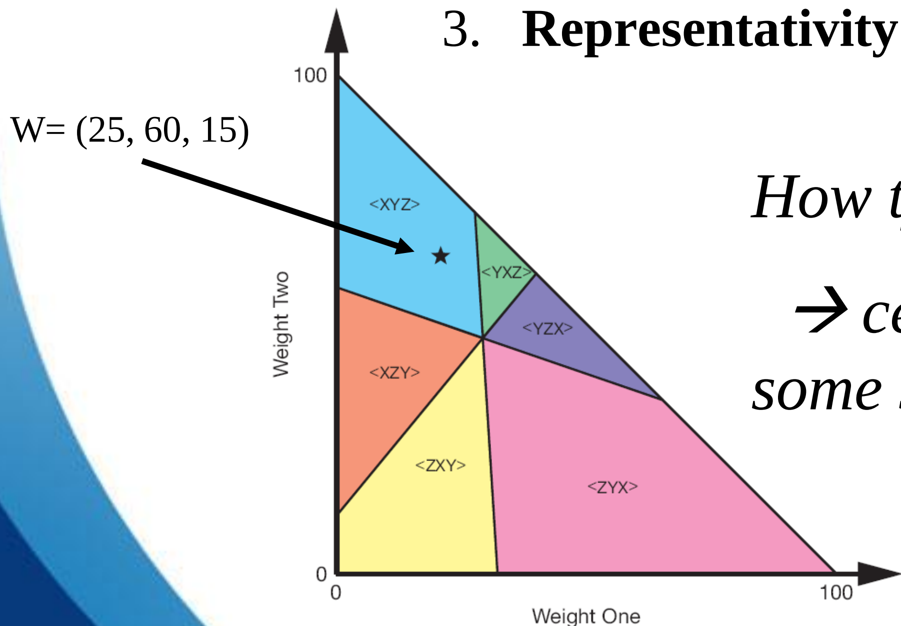
→ *probability of obtaining ranking when weights randomly selected*

Methodology

Geometric intuition:

Sensitivity measures:

1. Distance-based
2. Volume-based
3. **Representativity**



How typical are the chosen weights?

→ *central point of a region is in some sense most representative*

Methodology

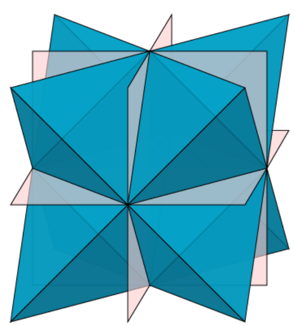
High-dimensional analogy

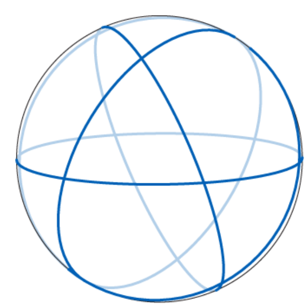
Hyperplane arrangements & polytopes:

$$\sum_{i=1}^{m} w_i = 100$$

$$w_i \geq 0 \quad \text{for } i = 1, \ldots, m$$

Score functions $S_j = \sum_{i=1}^{m} v_{ij} w_i$ and hyperplanes $\sum_{i=1}^{m} v_{ij} w_i - \sum_{i=1}^{m} v_{ik} w_i = 0$.



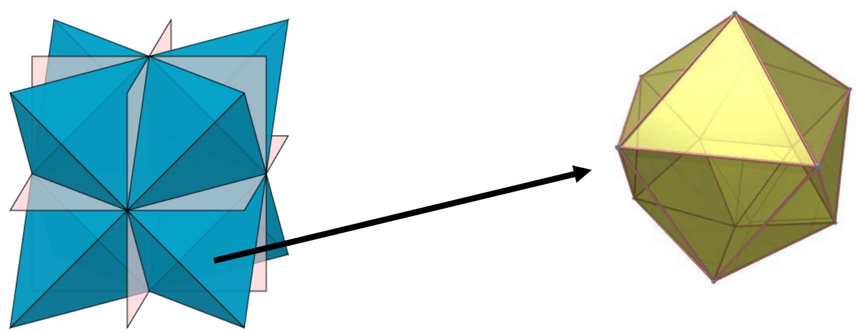


Methodology

High-dimensional analogy

Hyperplane arrangements & polytopes:

Each cell of hyperplane arrangement is an *(m-1)*-dim polytope **P**



$$P = \left\{ \vec{x} \in \mathbb{R}^d \;\middle|\; \sum_{j=1}^{d} a_{ij} x_j \leq b_i,\ i = 1, 2, \ldots, \bar{n} \right\}$$

Methodology

High-dimensional analogy

Hyperplane arrangements & polytopes:

Computational complexity:

The # of hyperplanes can be $\binom{n}{2}$

The number of polytopes (complete rankings of *n* options)

$$Q_n \approx \frac{K^{n+1}(n!)}{2} \quad \text{where} \quad K = \frac{1}{\ln 2} = 1.44269\ldots$$

Methodology:
Distance Sensitivity Measure

- *How close are baseline weights to nearest boundary?*

  → *to a weight being zero*

  → *to alternative ranking of options*

- Determine minimum required change to current weights to alter ranking of options

Method 1. Given a adjacent ranking polytope ***P*** and baseline weights ***W***:

find $\bar{W} = (\bar{W}_1, \bar{W}_2, \ldots, \bar{W}_m)^T \in P$ such that

$\boldsymbol{D} = \sqrt{\sum_{i=1}^{m} (\bar{W}_i - W_i)^2}$ is minimized (Quadratic Program)

Methodology:
Distance Sensitivity Measure

- *How close are baseline weights to nearest boundary?*
  → *to a weight being zero*
  → *to alternative ranking of options*
- Determine minimum required change to current weights to alter ranking of options

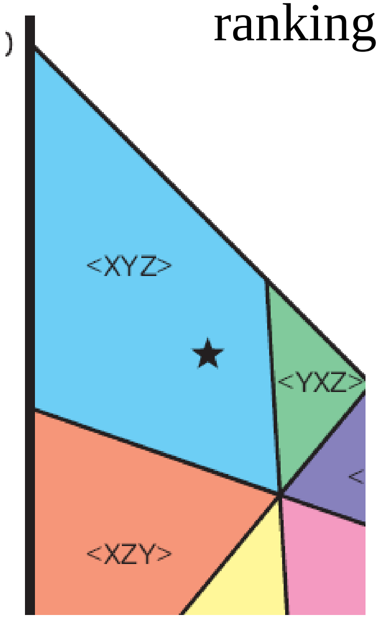
Method 2. Let ***P*** be polytope that contains baseline weights ***W***

$$P = \left\{ \vec{x} \in \mathbb{R}^d \;\middle|\; \sum_{j=1}^{d} a_{ij} x_j \leq b_i,\; i = 1, 2, \ldots, \bar{n} \right\}$$

Compute $\boldsymbol{D} = \min_i \left\{ b_i - \frac{\sum_{j=1}^{d} a_{ij} \boldsymbol{w}_j}{\sqrt{\sum_{j=1}^{d} a_{ij}^2}} \right\}$

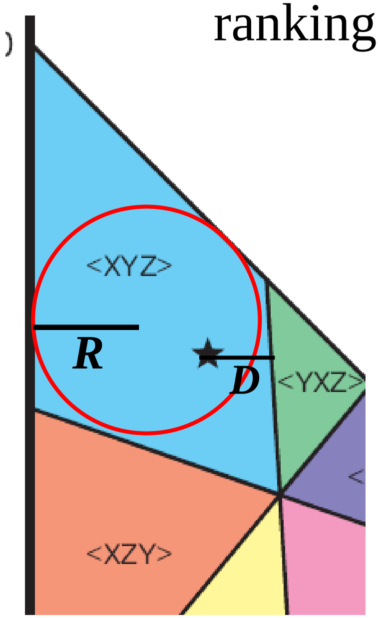
Methodology:
Distance Sensitivity Measure

- Computed ***D*** is only a relative measure: What to compare to?
- What is largest possible distance of any weights in the same ranking region?



Methodology: Distance Sensitivity Measure

- Computed ***D*** is only a relative measure: What to compare to?
- What is largest possible distance of any weights in the same ranking region?



» Compute radius of **Chebyshev Sphere** of the ranking region (polytope)

$$P = \left\{ \vec{x} \in \mathbb{R}^d \;\middle|\; \sum_{j=1}^{d} a_{ij} x_j \leq b_i,\ i = 1, 2, \ldots, \bar{n} \right\}$$

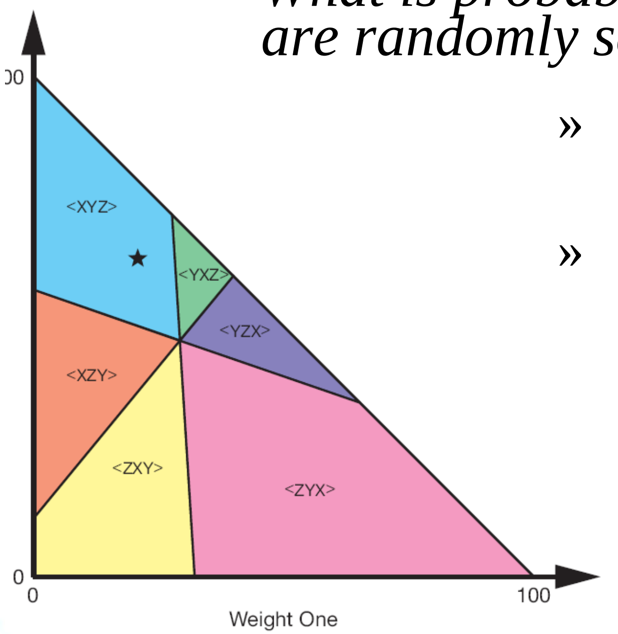
$$\text{maximize } r$$

$$\text{subject to } \sum_{j=1}^{d} a_{ij} x_j + r \sqrt{\sum_{j=1}^{d} a_{ij}^2} \leq b_i \quad \text{for all } i = 1, \ldots, \bar{n}$$

Compare ***D*** to ***R***: If $\frac{D}{R} \leq 0.05$ raise flag!

Methodology:
Volume Sensitivity Measure

- *How typical (large) is the ranking region?*
- *What is probability of obtaining ranking when weights are randomly selected?*
  - Compute volume of each ranking region and compare to volume of entire region
  - **Polytope volume computation is difficult**, but excellent codes exist for practical-sized instances (VINCI)



Let V = total volume of weight space

Let V^P = volume of ranking region P

$$\frac{V^P}{V} \leq \ ?$$

Methodology: Volume Sensitivity Measure

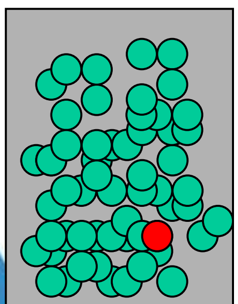
Let $\boldsymbol{V}$ = total volume of weight space

Let $\boldsymbol{V^P}$ = volume of ranking region $\boldsymbol{P}$

Claim: *If* $\frac{V^P}{V} \leq \frac{1}{3n!}$ *then raise flag!*

n! = *# of possible ranking regions*

Consider the case when all *n!* regions equally likely:



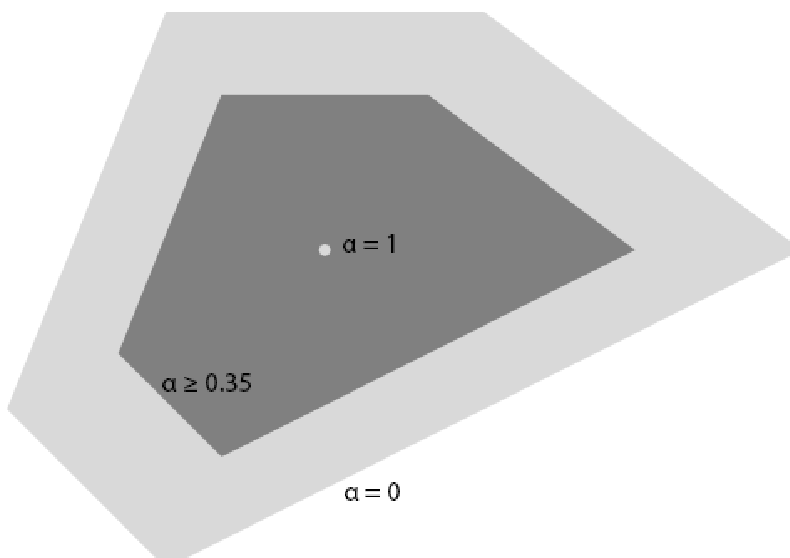
← Each ball represents a region in weight space.

How many draws (with replacement) until we are 95% certain to have drawn a specific (red) one?

» = 3*n!*

Methodology: Representativity Sensitivity Measure

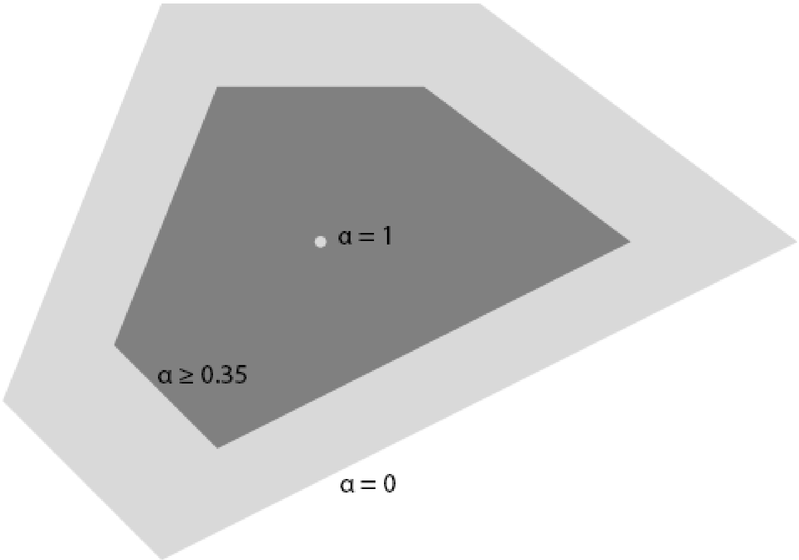
- *How typical are the chosen weights?*
- *Idea: central point of a region is in some sense most representative*



Methodology: Representativity Sensitivity Measure

- *How typical are the chosen weights?*
- *Idea: central point of a region is in some sense most representative*

Many definitions of a polytope centre…



Methodology: Representativity Sensitivity Measure

- *How typical are the chosen weights?*
- *Idea: central point of a region is in some sense most representative*

Suggest: Generalized barycentre (average of vertices)

$\boldsymbol{P} = \{\boldsymbol{x} \mid \boldsymbol{Ax} \leq \boldsymbol{b}\}$ and baseline weights $\boldsymbol{W}$

Let $\boldsymbol{C}$ = barycentre of $\boldsymbol{P}$

Find max $\boldsymbol{\alpha}$: $\boldsymbol{A(W\text{-}C)} \leq \boldsymbol{(1\text{-}\alpha)(b\text{-}AC)}$ with $\boldsymbol{0 \leq \alpha \leq 1}$

$\boldsymbol{\alpha}$ is representativity sensitivity measure

If $\boldsymbol{\alpha} \leq 0.05$ then raise flag!

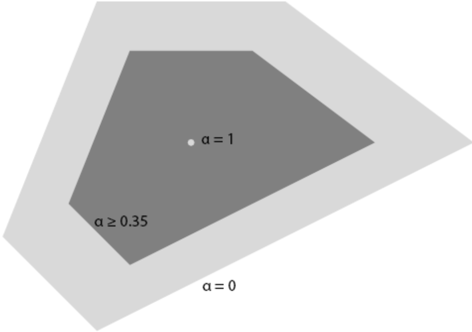
Methodology: Representativity Sensitivity Measure

Representativity relates to volume:

Claim: Given an $\boldsymbol{\alpha}$, the region defined by $\boldsymbol{Ax \leq (1-\alpha)b}$ has a volume that is

$(1 - \boldsymbol{\alpha})^{m-1}$ times the volume of $\boldsymbol{Ax \leq b}$

Proof: by triangulation of polytopes into simplices…



Example

Canadian Government needs to upgrade the capability of a class of warships

Options:

A: Refit current (old) warships

B: Buy existing warships from foreign country

C: Purchase foreign design and build in Canada

D: Design and build in Canada

Example

Canadian Government needs to upgrade the capability of a class of warships

7 Criteria:

1. In service support costs
2. Economic benefits
3. Sail-away costs
4. Operations & Doctrine
5. Schedule
6. Infrastructure requirements
7. Risk

Example

Criterion	Weight	A	B	C	D
		Option Ratings			
ISS costs	**50**	1	2	3	4
Economic benefits	**9**	2	1	3	4
Sail-away cost	**15**	4	2	3	1
Operations	**7**	3	1	2	4
Schedule	**10**	4	3	2	1
Infrastructure	**2**	4	2	2	2
Risk	**7**	2	3	1	4
		211	201	267	321

Final ranking is *< D C A B >*

How sensitive is the final ranking to changes in these weights?

Has someone fine-tuned the weights?

Example

Scoring functions:

$$
\begin{aligned}
S_D &= 4w_1 + 4w_2 + w_3 + 4w_4 + w_5 + 2w_6 + 4w_7,\\
S_C &= 3w_1 + 3w_2 + 3w_3 + 2w_4 + 2w_5 + 2w_6 + w_7,\\
S_A &= w_1 + 2w_2 + 4w_3 + 3w_4 + 4w_5 + 4w_6 + 2w_7,\\
S_B &= 2w_1 + w_2 + 2w_3 + w_4 + 3w_5 + 2w_6 + 3w_7.
\end{aligned}
$$

Polytope for ranking region < D C A B>

$$
\begin{aligned}
w_1 + w_2 + w_3 + w_4 + w_5 + w_6 + w_7 &= 100\\
w_1 + w_2 - 2w_3 + 2w_4 - w_5 + 3w_7 &\geq 0\\
2w_1 + w_2 - w_3 - w_4 - 2w_5 - 2w_6 - w_7 &\geq 0\\
-w_1 + w_2 + 2w_3 + 2w_4 + w_5 + 2w_6 - w_7 &\geq 0\\
w_i &\geq 0 \quad \text{for } i = 1, \ldots, 7.
\end{aligned}
$$

Example: Distance

Ranking	Distance	Weights						
		$\bar{W}_1$	$\bar{W}_2$	$\bar{W}_3$	$\bar{W}_4$	$\bar{W}_5$	$\bar{W}_6$	$\bar{W}_7$
$\langle DCAB \rangle$	0	50	9	15	7	10	2	7
$\langle DCBA \rangle$	3.035	51.71	8.87	13.95	5.95	9.87	0.95	8.70
$\langle CDAB \rangle$	12.84	48.54	7.54	23.0	2.38	14.85	3.69	0
$\langle DACB \rangle$	15.12	39.5	2.58	16.75	8.75	15.83	7.83	8.75
$\langle CADB \rangle$	17.39	41.75	3.50	23.25	4.25	18.25	7.5	1.5
$\langle ACDB \rangle$ $\langle ADCB \rangle$	17.61	40.34	2.72	22.18	5.27	18.34	8.11	3.04
$\langle DBCA \rangle$	19.56	45.50	0	10.5	2.5	17.50	3.50	20.50
$\langle CDBA \rangle$	20.02	60.71	0	21.43	0	17.85	0	0
$\langle DABC \rangle$ $\langle DBAC \rangle$	20.24	41.40	0	11.67	3.67	18.67	5.55	19.03

Radius of Chebyshev sphere of polytope < D C B A > = 9.38

test: 3.03 / 9.38 > 0.05

Example: Volume

Final Ranking	Fraction of Volume
$\langle DACB \rangle$	0.2637
$\langle ADCB \rangle$	0.2273
$\langle DCAB \rangle$	0.1125
$\langle DABC \rangle$	0.0961
$\langle ADBC \rangle$	0.0824
$\langle ACDB \rangle$	0.0820
$\langle ACBD \rangle$	0.0412
$\langle ABCD \rangle$	0.0310
$\langle ABDC \rangle$	0.0241
$\langle DBAC \rangle$	0.0180
$\langle DCBA \rangle$	0.0159
$\langle DBCA \rangle$	0.0031
$\langle CDAB \rangle$	0.0014
$\langle CADB \rangle$	0.0014
$\langle BADC \rangle$	1.1×10^{-5}
$\langle BDAC \rangle$	7.3×10^{-6}

***n* = 4**

test: 1 / (3n!) = 1.39% < 11.25%

Winner	Fraction
D	0.5092
A	0.4880
C	0.0028
B	1.8×10^{-5}

Example: Representativity

Baseline	Centroid of <D C A B>
50	26.33
9	29.01
15	14.64
7	9.60
10	4.49
2	8.59
7	7.34

Compute $\boldsymbol{\alpha}$:

test: $\boldsymbol{\alpha} = 0.1527 > 0.05$

The polytope with $\boldsymbol{\alpha} \geq 0.1527$ accounts for 37% of the original volume

Conclusion

- Weighted-sum methods very popular and easy to use except often require subjective choice of weights
- Using high-dimensional geometry one can analyze the sensitivity of the chosen weights
- We defined thresholds that can be used by decision makers to accept or reject weights

- Implementation available from authors (*Mathematica* Notebook)
- Publications:
  - *B. Kaluzny, R.H.A.D. Shaw, Sensitivity Analysis of Additive Weighted Scoring Methods, DRDC-CORA-TR-2009-002.*
  - *Submitted to Journal of Decision Analysis*

DEFENCE R&D DÉFENSE