**Distributionally Robust Optimization**

Congjian Wang and Diego Mandelli

May 13, 2020

# Stochastic Optimization Model Formulation for Capital Budgeting

The stochastic capital budgeting model is specified below:

*Indices and sets:*

|  |  |
| --- | --- |
|  | candidate projects |
|  | options for selecting project (e.g., initiate project in year or and in a standard (three year) or in an expedited (two year) manner) |
|  | must-do projects (e.g., due to safety reasons even if their NPV is negative) |
|  | option for project can be selected only if option is selected for project , i.e., piggybacking |
|  | types of resources, e.g., capital funds, O&M funds, labor-hours, time during outage |
|  | time periods (years) |
|  | scenarios |

*Data:*

|  |  |
| --- | --- |
|  | NPV (revenue less financial cost) of selecting project via option under scenario |
|  | available budget for a resource of type in year under scenario |
|  | consumption of resource of type in year if project is performed via option under scenario |
|  | probability mass of scenario |

*Decision variables:*

***Model formulation:***

|  |  |
| --- | --- |
|  | (1a) |
|  | (1b) |
|  | (1c) |
|  | (1d) |
|  | (1e) |
|  | (1f) |
|  | (1g) |
|  | (1h) |
|  | (1i) |
|  | (1j) |
|  | (1k) |
|  | (1l) |
|  | (1m) |

# Distributionally Robust Optimization Model Formulation for Capital Budgeting

A distributionally robust optimization variant of model (1) is then given by:

|  |  |
| --- | --- |
|  | (2a) |
|  | (2b) |
|  | (2c) |
|  | (2d) |
|  | (2e) |
|  | (2f) |
|  | (2g) |
|  | (2h) |
|  | (2i) |
|  | (2j) |
|  | (2k) |
|  | (2l) |
|  | (2m) |
|  | (2n) |
|  |  |

First we select , and then knowing , nature selects a worst-case probability distribution, , to minimize the expected NPV. If we choose to use Wasserstein distance to measure the distance between two probability distributions, i.e. a given candidate robust distribution , and a given distribution, . The constraint (2n) can be extended with given **radius of ambiguity**, (i.e. ) to:

|  |  |
| --- | --- |
|  | (2n-1) |
|  | (2n-2) |
|  | (2n-3) |
|  | (2n-4) |

Where is the general -norm distance between vector and , supposing is a discrete random variable. In capital budgeting, could be the realizations of available capital budgets.

## Reformulation via Duality Theory

For given , (i.e. fixed value for ), the distributionally robust optimization model (2) will be become:

|  |  |
| --- | --- |
|  | (3a) |
|  | (3b) |
|  | (3c) |
|  | (3d) |
|  | (3e) |

Define the *Lagrangian* function using nonnegative multipliers variables by taking the constraints out of the body of the model, and placing them in the objective function.

We then solve the presumably easier problem:

Compute the gradients, i.e.:

Let

One can obtain:

Since , we can extend constraint to . In this case, the optimization problem will reach its minimum at . Thus, the following constraints will be hold for the dual problem:

Or

The function is called the dual function, and we presume that computing is an easy task. The dual problem is then defined to be:

With constraint:

Thus the original problem will be reformulated as follows:

|  |  |
| --- | --- |
|  | (4a) |
|  | (4b) |
|  | (4c) |
|  | (4d) |
|  | (4e) |
|  | (4f) |
|  | (4g) |
|  | (4h) |
|  | (4i) |
|  | (4j) |
|  | (4k) |
|  | (4l) |
|  | (4m) |
|  | (4n) |
|  | (4o) |

# Distributionally Robust Optimization for Other Types of Capital Budgeting

The notation and formulation of the risk-informed models are as follows:

Indices and Sets:

|  |  |
| --- | --- |
|  | candidate projects |
|  | options for selecting project , e.g., initiate project in year or and in a standard  (three year) or in an expedited (two year) manner. Note that the last option for project is always used to indicate “non-selection”, i.e. the investment is not selected. |
|  | types of resources, e.g., capital funds, O&M funds, labor-hours, time during outage |
|  | time periods (years) |
|  | scenarios |

*Data:*

|  |  |
| --- | --- |
|  | profit of investment under scenario (NPV) |
|  | profit of investment via option under scenario (NPV) |
|  | available budget under scenario  |
|  | available budget for a resource of type under scenario  |
|  | available budget for unit under scenario  |
|  | available budget in year under scenario  |
|  | available budget for a resource of type in year under scenario  |
|  | cost of investment under scenario  |
|  | consumption of resource of type if investment is selected under scenario  |
|  | consumption of resource in year if investment is performed via option under scenario  |
|  | consumption of resource of type in year if investment is performed via option under scenario  |
|  | probability of scenario  |

*Decision variables:*



* 1. DRO for Single Knapsack Problem

|  |  |
| --- | --- |
|  | () |
|  | () |
|  | () |
|  | () |
|  | () |
|  | () |
|  | () |

* 1. DRO for Multi-Dimensional Knapsack Problem

|  |  |
| --- | --- |
|  | () |
|  | () |
|  | () |
|  | () |
|  | () |
|  | () |
|  | () |

* 1. DRO for Multiple Knapsack Problem

|  |  |
| --- | --- |
|  | () |
|  | () |

|  |  |
| --- | --- |
|  | () |

|  |  |
| --- | --- |
|  | () |

|  |  |
| --- | --- |
|  | () |

|  |  |
| --- | --- |
|  | () |

|  |  |
| --- | --- |
|  | () |
|  | () |

* 1. DRO for Multiple-Choice Knapsack Problem

|  |  |
| --- | --- |
|  | () |
|  | () |

|  |  |
| --- | --- |
|  | () |

|  |  |
| --- | --- |
|  | () |

|  |  |
| --- | --- |
|  | () |

|  |  |
| --- | --- |
|  | () |

|  |  |
| --- | --- |
|  | () |
|  | () |

# Appendix:

***Wasserstein Distance:***

|  |  |
| --- | --- |
|  | (5-a) |
|  | (5-b) |
|  | (5-c) |
|  | (5-d) |

Where is the general -norm distance between vector and , supposing is a discrete random variable. In capital budgeting, could be the realizations of available capital budgets. For any given radius of ambiguity, (i.e. ), formulation (2) will be extended as follows:

|  |  |
| --- | --- |
|  | (6-a) |
|  | (6-b) |
|  | (6-c) |
|  | (6-d) |

# Questions

1. How to involve distances for multiple parameters variations?
2. Does the dual solution provide the lower bound? Should we recompute the objective function with the solution of the dual problem?
3. Distance metric is dependent on the magnitude of the parameter values, it will have significant effects on the selection of .