May 4, 2022

It involves **maximizing the expected maximum** of a portfolio of random variables subject to a budget constraint.

Methodological orientation:

- Investment with uncertain payoff, search for the efficient frontier recall **portfolio allocation** models.
- Generalizes the **knapsack problem**: Integral packing constraint, NP-completeness, approximation algorithms.
- Objective is a **submodular set function**.

Today's presentation: **define the problem** and briefly summarize our **solution algorithms**. :

Market contains m **schools**, indexed by $\mathcal{C} = \{1 \dots m\}$. School j is named c_j.

We know the student's **admissions probability** f_j at each school.

Let the independent **random variable** $Z_j \sim \text{Bernoulli}(f_j) = 1$ if student is admitted, 0 otherwise.

Let $\mathcal{X} \subseteq \mathcal{C}$ denote the set of schools, or **application portfolio**, to which a student applies.

Application fees, time to write essays, and/or legal limits **constrain** applicant behavior. We consider a single knapsack constraint

$\sum_{j\in\mathcal{X}} g_j \leq H$ where g_j is called c_j's **application cost**. m , $\mathcal{C} = \{1 \dots m\}$. j c_j. This happens if and only if she **applies** to c_j, is **admitted** to c_j, and is **not admitted** to any school she prefers to c_j:

$$p_j(\mathcal{X}) = \begin{cases} f_j \prod_{\substack{i\in\mathcal{X}:\\ i>j}} (1 - f_i), & j \in \{0\} \cup \mathcal{X} \\ 0, & \text{otherwise.} \end{cases}$$

Therefore,

$$v(\mathcal{X}) = \sum_{j=1}^{m} t_j p_j(\mathcal{X}) = \sum_{j\in\mathcal{X}} \Big(f_j t_j \prod_{\substack{i\in\mathcal{X}:\\ i>j}} (1 - f_i) \Big).$$

[The college application problem, INLP form][,]
Another intuitive idea is the **linearization heuristic**: Since the expected utility associated with applying to c_j (alone) is $f_j t_j$, solve the knapsack problem

$$\text{maximize} \quad \sum_{j\in\mathcal{X}} f_j t_j \qquad \text{subject to} \quad \sum_{j\in\mathcal{X}} g_j \leq H$$

as a surrogate. This solution can be **arbitrarily bad**.

Fu (2014) solved a similar problem by **enumeration**, which is intractable for $m \geq 20$ or so.

Our algorithms provide both **time and accuracy guarantees**. Fu (2014) **n the "textbf special case where each**

$\backslash beamer@usebeamercolormathtextinlined\backslash ifbeamercolorempty[fg]math$

1**, we provide an**

$\backslash beamer@usebeamercolormathtextinlined\backslash ifbeamercolorempty[fg]math$

algorithm. , $m \geq 20$.

Solutions to the college application problem have **practical value**: US admissions consultants charge an average of \$200/hr! open-sourcing our code for public benefit. Lots of extensions to consider: parametric risk aversion, distribution constraints, FPTAS memory-usage improvements.

. 200! open source license .