Prophet Inequalities with Linear Correlations

Nicole Immorlica	Microsoft Research
Sahil Singla	Princeton
Bo Waggoner	**U. Colorado**

Oct. 2019

Outline:

1. Prophet inequalities - overview
2. This work: introducing correlations

A stopping time problem

Given: **independent** $X_1, \ldots, X_n$, known distributions

A stopping time problem

Given: **independent** $X_1, \ldots, X_n$, known distributions

Realizations are revealed one by one ...
... algorithm can stop at any time i and take X_i

A stopping time problem

Given: **independent** $X_1, \ldots, X_n$, known distributions

Realizations are revealed one by one ...
... algorithm can stop at any time i and take X_i

$\mathrm{OPT} := \max_i X_i$ *achieved by the prophet*

A stopping time problem

Given: **independent** $X_1, \ldots, X_n$, known distributions

Realizations are revealed one by one . . .
. . . algorithm can stop at any time i and take X_i

$\mathrm{OPT} := \max_i X_i$ *achieved by the prophet*

Goal: $\mathrm{ALG} \geq (?) \cdot \mathrm{OPT}$ *"prophet inequality"*

Known results

Optimal, backward-induction solution: $\mathrm{ALG} \geq 0.5 \cdot \mathrm{OPT}$.

[1] Further reading:
`http://bowaggoner.com/blog/2018/08-25-prophet-inequalities/index.html`

Known results

Optimal, backward-induction solution: $\mathrm{ALG} \geq 0.5 \cdot \mathrm{OPT}$.

Samuel-Cahn 1984: a **threshold policy** achieves 0.5:

1. Let $\tau =$ median of $\max_i X_i$
2. Stop at first X_i exceeding τ

[1] Further reading:
http://bowaggoner.com/blog/2018/08-25-prophet-inequalities/index.html

Known results

Optimal, backward-induction solution: $\mathrm{ALG} \geq 0.5 \cdot \mathrm{OPT}$.

Samuel-Cahn 1984: a **threshold policy** achieves 0.5:

1. Let $\tau =$ median of $\max_i X_i$
2. Stop at first X_i exceeding τ

Observed by Kleinberg+Weinberg 2012: $\tau = 0.5\,\mathbb{E}\left[\max_i X_i\right]$ also achieves 0.5 approximation ratio.[1]

[1] Further reading: `http://bowaggoner.com/blog/2018/08-25-prophet-inequalities/index.html`

Half-the-expected-max policy

Proof.

Let $P = \Pr[\max_i X_i \geq \tau]$.

$$\begin{aligned}
\mathbb{E}[\mathrm{ALG}] &= P \cdot \tau + \sum_{i=1}^{n} \Pr\left[\max_{i'<i} X_{i'} < \tau\right] \mathbb{E}\left[(X_i - \tau)^+\right] \\
&\geq P \cdot \tau + (1-P) \sum_{i=1}^{n} \mathbb{E}\left[(X_i - \tau)^+\right] \\
&\geq P \cdot \tau + (1-P)\, \mathbb{E}\left[\max_i (X_i - \tau)^+\right] \\
&\geq P \cdot \tau + (1-P) \left(\mathbb{E}\left[\max_i X_i\right] - \tau\right) \\
&\geq P \cdot \tau + (1-P)\tau \\
&= \tau.
\end{aligned}$$

Why threshold policies?

1. Threshold policies are **robust**
   Variables can arrive in any order, . . .

Why threshold policies?

1. Threshold policies are **robust**

   Variables can arrive in any order, ...

2. Single-item auction:

   - Buyers arrive sequentially with **secret** valuations X_i

Why threshold policies?

1. Threshold policies are **robust**
   Variables can arrive in any order, . . .
2. Single-item auction:
   - Buyers arrive sequentially with **secret** valuations X_i
   - Post a price τ
   - First buyer with $X_i \geq \tau$ purchases
   - "welfare" ≥ 0.5 optimal

This work: correlations

What if $X_1, \ldots, X_n$ are **correlated**?

(Known: constant-factor approx **cannot** be achieved)

This work: correlations

What if $X_1, \ldots, X_n$ are **correlated**?

(Known: constant-factor approx **cannot** be achieved)

Question 1: how to **model (limited) correlation**?

This work: correlations

What if $X_1, \ldots, X_n$ are **correlated**?

(Known: constant-factor approx **cannot** be achieved)

Question 1: how to **model (limited) correlation**?

Question 2: do **threshold policies** give prophet inequalities?

Outline:

- The linear correlations model
- Lower bound instance
- Key tool: Augmentation Lemma
- Results

Linear correlations model

Assume: there exist independent $Y_1, \ldots, Y_m$ such that

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{Y}$$

for $A \in \mathbb{R}_{\geq 0}^{m \times n}$.

Linear correlations model

Assume: there exist independent $Y_1, \ldots, Y_m$ such that

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{Y}$$

for $A \in \mathbb{R}_{\geq 0}^{m \times n}$.

Parameters:

- ℓ column sparsity (max. nonzero entries per column)
- k row sparsity

Linear correlations model

Assume: there exist independent $Y_1, \ldots, Y_m$ such that

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{Y}$$

for $A \in \mathbb{R}_{\geq 0}^{m \times n}$.

Parameters:

- ℓ column sparsity (max. nonzero entries per column)
- k row sparsity

Recall: Algorithm knows $\mathbf{A}$ and distributions of $\mathbf{Y}$, but only observes realizations of $\mathbf{X}$.

Lower bound

Theorem

No algorithm can guarantee better than $O\left(\frac{1}{\min\{\ell,k\}}\right)$ *approximation ratio.*

Lower bound

Theorem

No algorithm can guarantee better than $O\left(\frac{1}{\min\{\ell,k\}}\right)$ *approximation ratio.*

Tower variables:

$$Y_j = \begin{cases} \frac{1}{\epsilon^j} & \text{w.prob. } \epsilon^j \\ 0 & \text{o.w.} \end{cases}$$

$$\begin{aligned} X_1 &= Y_1 + \epsilon \cdot Y_2 + \cdots + \epsilon^n Y_n \\ X_2 &= Y_2 + \epsilon \cdot Y_3 + \cdots + \epsilon^{n-1} Y_n \\ &\vdots \\ X_{n-1} &= Y_{n-1} + \epsilon Y_n \\ X_n &= Y_n \end{aligned}$$

Relevent ideas.

Lower bound

Theorem

No algorithm can guarantee better than $O\left(\frac{1}{\min\{\ell,k\}}\right)$ *approximation ratio.*

Tower variables:

$$Y_j = \begin{cases} \frac{1}{\epsilon^j} & \text{w.prob. } \epsilon^j \\ 0 & \text{o.w.} \end{cases}$$

$$\begin{aligned} X_1 &= Y_1 + \epsilon \cdot Y_2 + \cdots + \epsilon^n Y_n \\ X_2 &= Y_2 + \epsilon \cdot Y_3 + \cdots + \epsilon^{n-1} Y_n \\ &\vdots \\ X_{n-1} &= Y_{n-1} + \epsilon Y_n \\ X_n &= Y_n \end{aligned}$$

Relevent ideas. Essentially, at most one Y_j is nonzero.

Lower bound

Theorem

No algorithm can guarantee better than $O\left(\frac{1}{\min\{\ell,k\}}\right)$ *approximation ratio.*

Tower variables:

$$Y_j = \begin{cases} \frac{1}{\epsilon^j} & \text{w.prob. } \epsilon^j \\ 0 & \text{o.w.} \end{cases}$$

$$\begin{aligned} X_1 &= Y_1 + \epsilon \cdot Y_2 + \cdots + \epsilon^n Y_n \\ X_2 &= Y_2 + \epsilon \cdot Y_3 + \cdots + \epsilon^{n-1} Y_n \\ &\vdots \\ X_{n-1} &= Y_{n-1} + \epsilon Y_n \\ X_n &= Y_n \end{aligned}$$

Relevent ideas. Essentially, at most one Y_j is nonzero.

$\mathrm{OPT} \approx n$ because it gets 1 from each Y_j on avg.

$\mathrm{ALG} \approx 1$ (consider dilemma when $X_i \neq 0$).

Lower bound

Theorem

No algorithm can guarantee better than $O\left(\frac{1}{\min\{\ell,k\}}\right)$ *approximation ratio.*

Tower variables:

$$Y_j = \begin{cases} \frac{1}{\epsilon^j} & \text{w.prob. } \epsilon^j \\ 0 & \text{o.w.} \end{cases}$$

$$\begin{aligned} X_1 &= Y_1 + \epsilon \cdot Y_2 + \cdots + \epsilon^n Y_n \\ X_2 &= Y_2 + \epsilon \cdot Y_3 + \cdots + \epsilon^{n-1} Y_n \\ &\vdots \\ X_{n-1} &= Y_{n-1} + \epsilon Y_n \\ X_n &= Y_n \end{aligned}$$

Relevent ideas. Essentially, at most one Y_j is nonzero.

$\mathrm{OPT} \approx n$ because it gets 1 from each Y_j on avg.

$\mathrm{ALG} \approx 1$ (consider dilemma when $X_i \neq 0$).

Even for $\ell = k = 2$, threshold algorithms achieve at best $\frac{1}{n}$.

Upper bounds

Main idea: We can achieve a matching $\Omega\left(\frac{1}{\min\{\ell,k\}}\right)$ bound by proving:

Theorem

There is an **inclusion-threshold** *algorithm achieving* $\mathrm{ALG} \geq \frac{1}{2e}\frac{1}{\ell}\mathrm{OPT}$.

Theorem

There is an **inclusion-threshold** *algorithm achieving* $\mathrm{ALG} \geq \frac{1}{2e^3}\frac{1}{k}\mathrm{OPT}$.

Upper bounds

Main idea: We can achieve a matching $\Omega\left(\frac{1}{\min\{\ell,k\}}\right)$ bound by proving:

Theorem

There is an **inclusion-threshold** *algorithm achieving* $\mathrm{ALG} \geq \frac{1}{2e}\frac{1}{\ell}\mathrm{OPT}$.

Theorem

There is an **inclusion-threshold** *algorithm achieving* $\mathrm{ALG} \geq \frac{1}{2e^3}\frac{1}{k}\mathrm{OPT}$.

Inclusion-threshold: Commit to **discarding** certain X_i *in advance*; apply a threshold to the rest.

Upper bounds

Main idea: We can achieve a matching $\Omega\left(\frac{1}{\min\{\ell,k\}}\right)$ bound by proving:

Theorem

There is an **inclusion-threshold** *algorithm achieving* $\mathrm{ALG} \geq \frac{1}{2e}\frac{1}{\ell}\mathrm{OPT}$.

Theorem

There is an **inclusion-threshold** *algorithm achieving* $\mathrm{ALG} \geq \frac{1}{2e^3}\frac{1}{k}\mathrm{OPT}$.

Inclusion-threshold: Commit to **discarding** certain X_i *in advance*; apply a threshold to the rest.

Hope: approximate original independent prophets problem on a $\frac{1}{\ell}$ fraction of the input.

Problem: correlations still remain!

Key tool: Augmented Prophets Problem

Prophet instances of the form $X_i = Z_i + W_i$ where:

- Z_i is **independent** of $X_1, \ldots, X_{i-1}$

Key tool: Augmented Prophets Problem

Prophet instances of the form $X_i = Z_i + W_i$ where:

- Z_i is **independent** of $X_1, \ldots, X_{i-1}$
- **but** W_i may be chosen adversarially conditioned on $Z_1, W_1, \ldots, Z_i$.

Key tool: Augmented Prophets Problem

Prophet instances of the form $X_i = Z_i + W_i$ where:

- Z_i is **independent** of $X_1, \ldots, X_{i-1}$
- **but** W_i may be chosen adversarially conditioned on $Z_1, W_1, \ldots, Z_i$.
- **however** $\mathrm{OPT} = \max_i Z_i$.

Key tool: Augmented Prophets Problem

Prophet instances of the form $X_i = Z_i + W_i$ where:

- Z_i is **independent** of $X_1, \ldots, X_{i-1}$
- **but** W_i may be chosen adversarially conditioned on $Z_1, W_1, \ldots, Z_i$.
- **however** $\mathrm{OPT} = \mathrm{max}_i\, Z_i$.

Story:

- We have a standard prophets problem $Z_1, \ldots, Z_n$.
- But a **mischievious genie** intercepts and **augments** arrivals with W_i
- Genie can only increase X_i but tries to mess up ALG
- The genie cannot see the future (Z_{i+1} is independent of W_i, etc)

Key tool: Augmented Prophets Problem

Prophet instances of the form $X_i = Z_i + W_i$ where:

- Z_i is **independent** of $X_1, \ldots, X_{i-1}$
- **but** W_i may be chosen adversarially conditioned on $Z_1, W_1, \ldots, Z_i$.
- **however** $\mathrm{OPT} = \max_i Z_i$.

Story:

- We have a standard prophets problem $Z_1, \ldots, Z_n$.
- But a **mischievious genie** intercepts and **augments** arrivals with W_i
- Genie can only increase X_i but tries to mess up ALG
- The genie cannot see the future (Z_{i+1} is independent of W_i, etc)

Fact: median-of-max rule achieves 0 on augmented prophets problem!
$X_i =$ i.i.d. Bernoulli(ϵ); augment first arrival slightly.

Lemma (Augmentation Lemma)

Setting a threshold $\tau = 0.5\,\mathbb{E}\left[\max_i Z_i\right]$ *achieves* $\mathrm{ALG} \geq 0.5 \cdot \mathrm{OPT}$ *on the augmented prophets problem.* $\implies$ ignore the genie!

Proof.

Let $P = \Pr[\max_i X_i \geq \tau]$. Let E_i be event that $\max_{i'<i} X_{i'} < \tau$.

$$\begin{aligned}
\mathbb{E}[\mathrm{ALG}] &= P \cdot \tau + \sum_{i=1}^{n} \Pr\left[E_i\right] \mathbb{E}\left[(X_i - \tau)^+ \;\middle|\; E_i\right] \\
&\geq P \cdot \tau + (1-P) \sum_{i=1}^{n} \mathbb{E}\left[(Z_i - \tau)^+ \;\middle|\; E_i\right] \\
&\geq P \cdot \tau + (1-P) \sum_{i=1}^{n} \mathbb{E}\left[(Z_i - \tau)^+\right] \\
&= \tau.
\end{aligned}$$

$\square$

Proof of Theorem 2

Column sparsity ℓ: Each Y_j appears in at most ℓ different arrivals X_i.
Algorithm:

1. Include each X_i independently with prob. $\frac{1}{\ell}$; discard others.
2. Let $T_i = \{j : A_{ij} > 0$ and for all included $i' < i$, $A_{i'j} = 0\}$.
3. Let $Z_i = \sum_{j \in T_i} A_{ij} Y_j$
4. Let $W_i = \sum_{j \notin T_i} A_{ij} Y_j$
5. Solve augmented prophets problem on only included $X_i = Z_i + W_i$

Proof of Theorem 2

Column sparsity ℓ: Each Y_j appears in at most ℓ different arrivals X_i.
Algorithm:

1. Include each X_i independently with prob. $\frac{1}{\ell}$; discard others.
2. Let $T_i = \{j : A_{ij} > 0 \text{ and for all included } i' < i,\ A_{i'j} = 0\}$.
3. Let $Z_i = \sum_{j \in T_i} A_{ij} Y_j$
4. Let $W_i = \sum_{j \notin T_i} A_{ij} Y_j$
5. Solve augmented prophets problem on only included $X_i = Z_i + W_i$

Fact (Augmentation Lemma): $\mathrm{ALG} \geq \frac{1}{2}\,\mathbb{E}[\max_i Z_i]$.

Proof of Theorem 2

Column sparsity ℓ: Each Y_j appears in at most ℓ different arrivals X_i.
Algorithm:

1. Include each X_i independently with prob. $\frac{1}{\ell}$; discard others.
2. Let $T_i = \{j : A_{ij} > 0$ and for all included $i' < i$, $A_{i'j} = 0\}$.
3. Let $Z_i = \sum_{j \in T_i} A_{ij} Y_j$
4. Let $W_i = \sum_{j \notin T_i} A_{ij} Y_j$
5. Solve augmented prophets problem on only included $X_i = Z_i + W_i$

Fact (Augmentation Lemma): $\mathrm{ALG} \geq \frac{1}{2}\,\mathbb{E}[\max_i Z_i]$.

Claim: $\mathbb{E}[\max_i Z_i] \geq \frac{1}{e}\frac{1}{\ell}\,\mathbb{E}[\max_i X_i]$.
Proof: each Y_j appears in exactly one included X_i w.prob. $\geq \frac{1}{e}\frac{1}{\ell}$.

Proof of Theorem 3

Row sparsity k: Each X_i depends on at most k different variables Y_j.

Observation:

- Take any instance
- Prepend it with 1 million copies of $X_i = 0.00000001 Y_1$
- Row sparsity is unchanged

Proof of Theorem 3

Row sparsity k: Each X_i depends on at most k different variables Y_j.

Observation:

- Take any instance
- Prepend it with 1 million copies of $X_i = 0.00000001 Y_1$
- Row sparsity is unchanged

Conclusion: Subsampling must depend on "importance" of each row.

Proof of Theorem 3

Row sparsity k: Each X_i depends on at most k different variables Y_j.

Observation:

- Take any instance
- Prepend it with 1 million copies of $X_i = 0.00000001 Y_1$
- Row sparsity is unchanged

Conclusion: Subsampling must depend on "importance" of each row.

Algorithm outline:

1. Construct subsampled set S of arrivals *next!*
2. Construct Z_i, W_i and use augmentation threshold as before

Proof of Theorem 3

Row sparsity k: Each X_i depends on at most k different variables Y_j.

Observation:

- Take any instance
- Prepend it with 1 million copies of $X_i = 0.00000001 Y_1$
- Row sparsity is unchanged

Conclusion: Subsampling must depend on "importance" of each row.

Algorithm outline:

1. Construct subsampled set S of arrivals *next!*
2. Construct Z_i, W_i and use augmentation threshold as before

Claim: We have a subsampling scheme such that
$\mathbb{E}[\max_i Z_i] \geq \frac{1}{e^3} \frac{1}{k} \mathbb{E}[\max_i X_i]$.

Claim: We have a subsampling scheme such that
$\mathbb{E}[\max_i Z_i] \geq \frac{1}{e^3}\frac{1}{k}\,\mathbb{E}[\max_i X_i]$.

Scheme to construct $S \subseteq \{1, \ldots, n\}$:

1. For each Y_j, let $i^*(j) = \arg\max_i A_{ij}$.
2. Create graph on $\{1, \ldots, m\}$ with edge (j, j) if $A_{i^*(j)j'} > 0$.
3. Permute $\{1, \ldots, m\}$ such that for all t, there are at most k edges from vertices $\pi(1), \ldots, \pi(t-1)$ to $\pi(t)$.
4. For $t = 1, \ldots, m$, w.prob. $\frac{1}{k}$ add $i^*(\pi(t))$ to S and delete all vertices from π with edges to or from $\pi(t)$.

Claim: We have a subsampling scheme such that
$\mathbb{E}[\max_i Z_i] \geq \frac{1}{e^3}\frac{1}{k}\,\mathbb{E}[\max_i X_i]$.

Scheme to construct $S \subseteq \{1, \ldots, n\}$:

1. For each Y_j, let $i^*(j) = \arg\max_i A_{ij}$.
2. Create graph on $\{1, \ldots, m\}$ with edge (j, j) if $A_{i^*(j)j'} > 0$.
3. Permute $\{1, \ldots, m\}$ such that for all t, there are at most k edges from vertices $\pi(1), \ldots, \pi(t-1)$ to $\pi(t)$.
4. For $t = 1, \ldots, m$, w.prob. $\frac{1}{k}$ add $i^*(\pi(t))$ to S and delete all vertices from π with edges to or from $\pi(t)$.

Claim 1: Such a permutation π exists.

Out-degree $\leq k$, so average in-degree $\leq k$, so some vx can be placed last; repeat.

Claim: We have a subsampling scheme such that
$\mathbb{E}[\max_i Z_i] \geq \frac{1}{e^3}\frac{1}{k}\,\mathbb{E}[\max_i X_i]$.

Scheme to construct $S \subseteq \{1,\dots,n\}$:

1. For each Y_j, let $i^*(j) = \arg\max_i A_{ij}$.
2. Create graph on $\{1,\dots,m\}$ with edge (j,j) if $A_{i^*(j)j'} > 0$.
3. Permute $\{1,\dots,m\}$ such that for all t, there are at most k edges from vertices $\pi(1),\dots,\pi(t-1)$ to $\pi(t)$.
4. For $t = 1,\dots,m$, w.prob. $\frac{1}{k}$ add $i^*(\pi(t))$ to S and delete all vertices from π with edges to or from $\pi(t)$.

Claim 1: Such a permutation π exists.
Out-degree $\leq k$, so average in-degree $\leq k$, so some vx can be placed last; repeat.

Claim 2: $Z_{i^*(j)} = Y_j$ w.prob. $\geq \frac{1}{e^3}\frac{1}{\ell}$ (and then no $Z_{i'}$ includes Y_j).
It has $\leq 2k$ edges to earlier vertices, which all fail w.prob. $\geq \frac{1}{e^2}$; then it is chosen w.prob. $\frac{1}{\ell}$; then all others with $A_{i^*(j)j'} > 0$ fail to be included w.prob. $\geq \frac{1}{e}$.

Extensions

Suppose: algorithm can take $\leq r$ arrivals.

"r-uniform matroid constraint"

$\mathrm{OPT} =$ sum of r largest X_is

Extensions

Suppose: algorithm can take $\leq r$ arrivals.

"r-uniform matroid constraint"

$\mathrm{OPT} =$ sum of r largest X_is

Observation: Lower-bound can be extended: For fixed ℓ, unbounded row sparsity, no algorithm beats $O\left(\frac{1}{\ell}\right)$.

Extensions

Suppose: algorithm can take $\leq r$ arrivals.
"r-uniform matroid constraint"
$\mathrm{OPT} =$ sum of r largest X_is

Observation: Lower-bound can be extended: For fixed ℓ, unbounded row sparsity, no algorithm beats $O\left(\frac{1}{\ell}\right)$.

Theorem: For fixed k, as $n, r \to \infty$, we can achieve $1 - o(1)$ approximation ratio.

Extensions

Suppose: algorithm can take $\leq r$ arrivals.
"r-uniform matroid constraint"
$\mathrm{OPT} =$ sum of r largest X_is

Observation: Lower-bound can be extended: For fixed ℓ, unbounded row sparsity, no algorithm beats $O\left(\frac{1}{\ell}\right)$.

Theorem: For fixed k, as $n, r \to \infty$, we can achieve $1 - o(1)$ approximation ratio.

Key ingredient: An **Augmentation Lemma** for the cardinality-r prophet problem.
Much harder!

Recap

Prophet problem with **linear correlations:**

$$\mathbf{X} = \mathbf{A} \cdot \mathbf{Y}$$

Augmentation Lemma: There exists a 0.5-approx-ratio alg. for the augmented prophets problem.

Main result: Inclusion-threshold algorithms achieve

$$\Omega\left(\frac{1}{\min\{\text{row sparsity}, \text{col sparsity}\}}\right)$$

and this is tight for any algorithm.

Tight results for cardinality-k version as well; reveals unbounded col. sparsity is the harder problem.