Flexible Estimation to Sequential Search: A Partial Ranking Structure

Tinghan Zhang

Tilburg School of Economics and Management
Department of Econometrics and Operation Research

Sep, 2024

Motivation: Sequential Search Model

- Imagine that a consumer plans to purchase a product among many alternatives in a large market, but she has only partial information about each product.
- The consumer can spend time and attention to collect detailed product information sequentially to help her make better purchase decision.
- This is the basic setup of the Consumer Sequential Search Model (SSM).
- Weitzman (1979) proposes stepwise Optimal Search Rules to describe the optimal solution to SSM.
- Yet, the Optimal Search Rules are not empirically friendly.
  - The optimal decision in each step depends on unobserved search outcomes in previous steps.
  - It is not easy to decompose joint probability.
  - The estimation is either difficult in computation, lacking precision, or complicated in implementation.
  - Inflexible under partial/extra data and model variations.
- Either use the full model with a heavy implementation burden, or discard search information.

This paper

- This paper aids the simple and flexible empirical application of SSM.
- I propose four conditions equivalent to (yet do not rely on) Weitzman's rules. The conditions form a Partial Ranking (PR) structure.
- With the PR structure, the probability of observations can be decomposed to independent conditionals.
    - Also easy for specifying identification arguments.
- Flexible for full data, partial data, extra information, or tractable model variations.
    - Adaptive to the standard discrete choice structure on Choi et al.'s (2018) Eventual Purchase Theorem.
    - For more complicated variations, I provide an estimator with good performance for the search-with-product-discovery model.
    - The other example is my JMP, in which preference discovery alters the ranking in the middle of search.

Baseline Model: Sequential Search Model (SSM)

- A consumer i plans to purchase one product from a set of alternatives $\mathcal{M}_i$.
- The consumer has full knowledge of $\mathcal{M}_i$, but partial knowledge of each product in $\mathcal{M}_i$.
- The consumer can inspect products sequentially: she expends a search cost and fully resolves a product's uncertainty. Match value is determined once a product is inspected.
- The consumer can stop searching and buy one inspected product after each inspection.

- Data of each consumer: purchased product, set of inspected products, order of inspections, $\mathcal{M}_i$.
- $\{H, \mathcal{S}, \mathcal{R}, \mathcal{M}\}_i$: sequence observation of consumer i.
- Number inspected products following $\mathcal{R}_i$: $\{1, \cdots, J\}_{\mathcal{R}_i}$.
  Randomly number uninspected products with $J+1, \cdots, |\mathcal{M}_i|$.
- (Always) mark the number of the purchased product (H) by h. $1 \leq h \leq J$.

Baseline Model: Value of Inspection

- Assume search costs c_{ij} is independent, invariant, and observed by the consumer i.
- Weitzman (1979) simplifies consumers' dynamic optimization problem of sequential search model. He first introduced the value of an inspection.
- Imagine you have an alternative option that offers you a determined value of $\bar{u}$. Then inspecting an additional product j is indifferent when:

$$\underbrace{-c_{ij}}_{\text{Search cost}} + \underbrace{\int_{u>\bar{u}} (u - \bar{u})\, dF_{ij}^{u}(u)}_{\text{Expected extra gain}} = 0 \tag{1}$$

- Unique solution z_{ij}. Inspect j when $\bar{u}$ larger than z_{ij}; not inspect j when $\bar{u}$ is smaller than z_{ij}.
- z_{ij} is considered as the value of inspecting product j, or the reservation value of j.
- c_{ij} is only relevant to the model through z_{ij}.

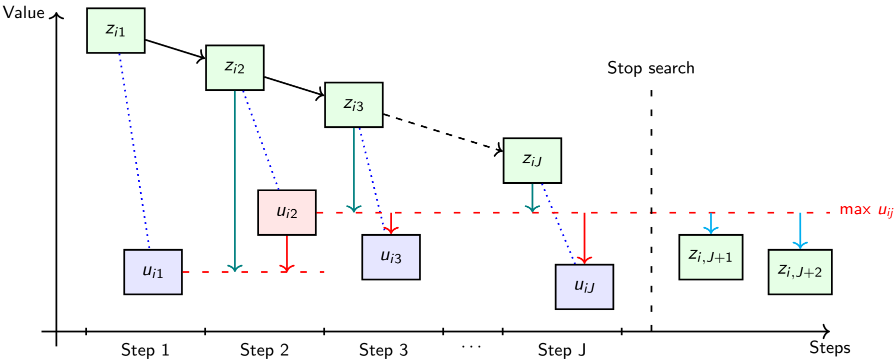
Optimal Search Rules (Weitamzn, 1979)

- "If a box is to be opened, it should be that closed box (*products not inspected*) with highest reservation price (*reservation value*)."
- "Terminate search whenever the maximum sampled reward (*match value*) exceeds the reservation price of every closed box."
- In empirical, we add one rule: "Select the opened box with the highest sampled reward. "

- Joint probability: all three rules hold.
- These rules are interdependent with unobserved search outcomes.

Optimal Search Rules (OSR) Structure

- Lead to the structure of Optimal Search Rules (OSR):



- Step-by-step structure. All solid arrows are supposed to hold.
- Interdependence: later choices are made conditional on outcomes from previous steps.
- Difficult to decompose probability, specify identification arguments, or implement estimation.

Partial Ranking (PR) Structure

Proposition 1

Define $y_i = \min\{u_{ih}, z_{iJ}\}$ *the Core Value of consumer* i*. Weitzman's optimal rules are fulfilled if and only if the following conditions are fulfilled:*

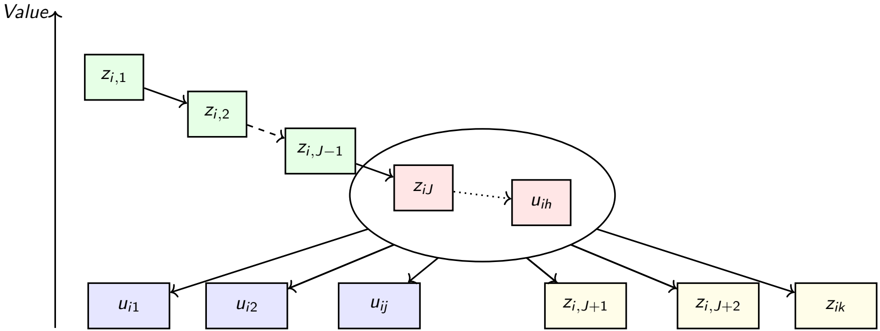
1. *Distribution Condition:* $u_{ih} \leq z_{iJ}$ *if* $h < J$*.*
2. *Ranking Condition:* $z_{i1} \geq z_{i2} \geq \ldots \geq z_{iJ}$*;*
3. *Choice Condition 1:* $z_{ik} \leq y_i$ *for all* $k > J$*;*
4. *Choice Condition 2:* $u_{ij} \leq y_i$ *for all* $j \leq J, j \neq h$*.*

- The joint probability:

$$\Pr(\{H, S, R, \mathcal{M}\}_i) = \Pr(z_{iJ} \geq u_{ih} \cap z_{i1} \geq \ldots \geq z_{iJ} \cap \max_{j \leq J} u_{ij} \leq y_i \cap \max_{k > J} z_{ik} \leq y_i)$$

Partial Ranking (PR) Structure: Illustration

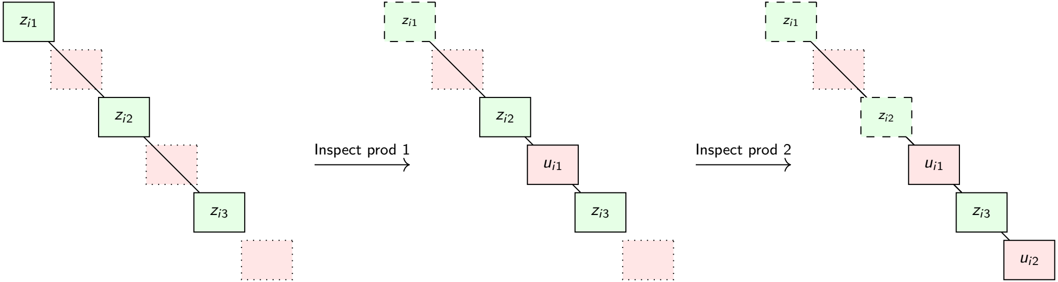
- The four conditions form the Partial Ranking (PR) structure, illustrated as follows:



- Static structure.
- The search process, as well as the eventually unpurchased and uninspected products, are only conditional on the core value.

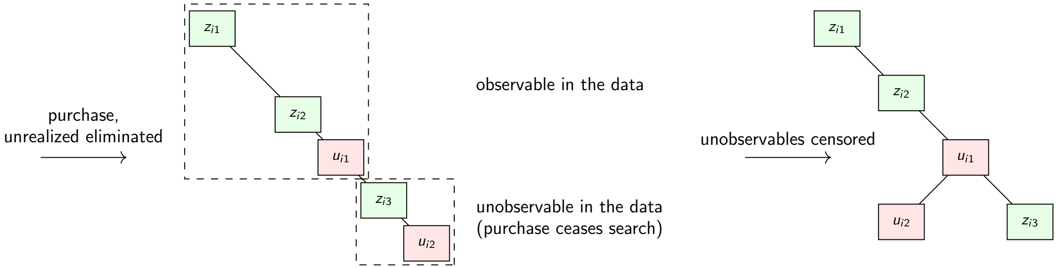
Partial Ranking (PR) Structure: Optimality

- The optimality of the PR structure does not rely on the Optimal Search Rules.
- Key idea: parameters (preferences, search costs) are fully informed by the ranking of MVs and RVs.
- The ranking remains stable but not fully revealed. Initially, consumers only observe RVs.
- Every inspection collapses an RV and reveals an MV without changing values.



Partial Ranking (PR) Structure: Optimality (Cont'd)

- Optimal: take actions following the descending order of the ranking (Keller & Oldale, 2003).
- Search stops when acting on an alternative with MV, i.e., purchase.
- At last, RVs of all and MVs of inspected are revealed; MVs of uninspected are eliminated.
- Part of the ranking is censored. Values of uninspected and unpurchased are smaller than y_i.



Partial Ranking (PR) Structure: Joint Probability

- Take the following value specification as an example (Honka and Chintagunta, 2017):

$$u_{ij} = X_i\gamma + p_{ij}\beta_i + \zeta_{ij} + \varepsilon_{ij}, \quad c_{ij} = c, \quad z_{ij} = X_i\gamma + p_{ij}\beta_i + \zeta_{ij} + m_\varepsilon(c).$$

  - It can be proved that z_{ij} follows a linear specification. $\delta(.)$ is derived from Equation (1).
  - ζ_{ij} is a pre-inspection taste shock for product j.

- Assumptions:
  1. Consumer knows $F^{\varepsilon}_{ij}(.) = F^{\varepsilon}(.)$, but not ε_{ij} until inspecting j.
  2. Consumer observes ζ_{ij} at the beginning of search.
  3. (Independence) Taking action on related products does not lead to information on other products.
  4. (Invariance) No external factor changes product values throughout the search process.

- Stack product values for vectorized representation:

$$\boldsymbol{z}_i^k = (z_{i,J}, \cdots, z_{i,1})^\top, \ \boldsymbol{z}_i^u = (z_{i,J+1}, \cdots, z_{i,|\mathcal{M}_i|})^\top, \ \boldsymbol{u}_i^{k'} = (u_{i,1}, \cdots, u_{i,h-1}, u_{i,h+1}, \cdots u_{i,J})^\top$$

Partial Ranking (PR) Structure: Joint Probability

- Joint Probability of SSM when $h < J$:

$$\Pr\left(\underbrace{\hat{D}}_{(J+|\mathcal{M}_i|-1)\times(J+|\mathcal{M}_i|)} \begin{pmatrix} u_{ih} \\ \boldsymbol{z}_i^k \\ \boldsymbol{z}_i^u \\ \boldsymbol{u}_i^{k'} \end{pmatrix}_{(J+|\mathcal{M}_i|)\times 1} \leq \boldsymbol{0} \right) = \Pr\left(\hat{D} \begin{pmatrix} \varepsilon_{ih} + \zeta_{ih} \\ \zeta_i^k \\ \zeta_i^u \\ \boldsymbol{\varepsilon}_i^{k'} + \zeta_i^{k'} \end{pmatrix} \leq -\hat{D} \begin{pmatrix} X_{ih}\gamma + p_{ih}\beta_i \\ \boldsymbol{X}_i^k\gamma + \boldsymbol{p}_i^k\beta_i + \vec{m}_\varepsilon(c) \\ \boldsymbol{X}_i^u\gamma + \boldsymbol{p}_i^u\beta_i + \vec{m}_\varepsilon(c) \\ \boldsymbol{X}_i^{k'}\gamma + \boldsymbol{p}_i^{k'}\beta_i \end{pmatrix} \right).$$

- The full-rank difference matrix $\hat{D} = \begin{pmatrix} \hat{D}_1 & \hat{D}_3 \\ \hat{D}_2 & \hat{D}_4 \end{pmatrix}$:

$$\hat{D}_1 = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & -1 & 0 \\ 0 & 0 & \cdots & 0 & 1 & -1 \end{pmatrix}_{J\times(J+1)}, \quad \hat{D}_2 = \begin{pmatrix} -1 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ -1 & 0 & \cdots & 0 \end{pmatrix}_{(|\mathcal{M}_i|-1)\times(J+1)},$$

$$\hat{D}_3 = \{0\}_{J\times(|\mathcal{M}_i|-1)}, \quad \hat{D}_4 = I_{(|\mathcal{M}_i|-1)\times(|\mathcal{M}_i|-1)}$$

- When $h = J$, the rank of the difference matrix $\tilde{D}$ is also $J + |\mathcal{M}| - 1$.

Partial Ranking (PR) Structure: Identification

- For either $D \in \{\hat{D}, \tilde{D}\}$, it does not differentiate model identification from an SDC model.
- The introduction of heterogeneity in RV is important for "Only difference matters."
- The standard deviation of δ_{ij} scales the model.
  - Notice that it scales $m_\varepsilon(c)$ but not c.
- σ_ε and c are two determinant of $m_\varepsilon(c)$. Identifying c requires previous identification of σ_ε.
- Identifying σ_ε from the choices is fragile because of the heteroskedasticity without exclusion restriction on correlations (Keane, 1992).
- Honka and Chintagunta (2017): Estimate c conditional on an extra assumption on $\sigma_\varepsilon = 1$.
  - The estimated search cost is very sensitive to the choice of σ_ε.

.

Partial Ranking (PR) Structure: Estimation

- Following the OSR structure, estimating SSM is practically difficult due to interdependency.

- The widely-applied simulator under OSR: Kernel-Smoothed Frequency Simulator.
  - Calculate $t_{ij}^1, t_{ij}^2, t_i^3, t_i^4$ for each observation i. Smooth with a kernel and scaling factors $\{\rho_1, \rho_2, \rho_3, \rho_4\}$.
  - Highly sensitive to the scaling factors. Needs pre-calibration on an artificial dataset.
  - More complicated model: more scaling factors. "Curse of dimensionality" for researchers.

- Recent development (Chung et al., 2024; Jiang et al., 2021): OSR-GHK simulator.
  - The simulator is smooth and efficient. No smoothing factors are needed.
  - Complicated in implementation: separate observations into 3 or 4 different cases before calculating the likelihood for each case.

Partial Ranking (PR) Structure: Estimation (Cont'd)

- PR-GHK simulator is flexible because each part of the joint probability is only related to the core value.
- One can easily adjust the structure and calculate the ranking.
- Compared to the KSFS: higher precision, circumventing pre-calibration on scaling factors.
- Compared to the OSR-GHK: almost the same efficiency, simpler implementation, higher flexibility.

Extension 1: Compatibility to Partial Data

Corollary 1

When a product is known to consumer (its match value is determined) without search, if it is not purchased, its match value follows Choice Condition 2; if it is purchased, all other products follow conditions in Proposition 1.

- Adding a known product (e.g., an outside option) does not affect the structure.
- Also when information on the inspection of some products is missing.
- If all products are known without inspection, Distribution and Rank Conditions are trivial.
  - With only two Choice Conditions, the PR structure degenerates to an SDC structure.

Extension 1: Compatibility to Partial Data

- When $\mathcal{S}_i$ or $\mathcal{R}_i$ is unavailable, summing up all possible $\mathcal{S}_i$ coincides with the SDCM based on the Eventual Purchase Theorem proposed by Choi et al. (2018).

Proposition 2 (when $\mathcal{S}_i$ is unavailable)

Define $w_{ij} = \min\{z_{ij}, u_{ij}\}$ the Effective Value of product j to consumer i. If $w_{iH} \geq w_{iL}, \forall L \in \mathcal{M}_i \backslash \{H\}$, then following Proposition 1, H is always inspected and purchased. On contrary, $w_{ih} \geq w_{ij}, \forall j \neq h$ must hold for any $\{H, \mathcal{S}, \mathcal{R}, \mathcal{M}\}_i$ fulfilling conditions in Proposition 1.

Corollary 2 (when $\mathcal{R}_i$ is unavailable)

A product H in $\mathcal{S}_i$ is purchased if and only if:

1. $u_{iL} < w_{iH} < z_{iL}, \forall L \in \mathcal{S}_i \backslash \{H\}$,
2. $w_{iH} > z_{iL'}, \forall L' \notin \mathcal{S}_i$

- More convenient for demand estimation, while information in the search process is left out.

Extension 1: Flexible Estimation

- PR-GHK simulator is flexible with partial data.
- When some data is missing, restructure the ranking condition over the missing values.
- Due to the independence between conditionals, no effect on the implementation of the other parts.
- Performs well when only the first inspection and the final purchase are observed:

	True value	PR-GHK Estimates
γ_1	1	0.999 (0.003)
γ_2	0.5	0.501 (0.002)
γ_3	-0.2	-0.202 (0.002)
β	-0.6	-0.608 (0.003)
σ_β	0.2	0.207 (0.003)
c	-1.5	-1.454 (0.011)

Extension 2: Variation on the Theme

- We can take additional information into the joint probability for estimation, as long as the ranking condition remains traced throughout the search process, including:
    - Extra information on the unobserved ranking (e.g., the 'second choice').
    - Unforeseen shocks that vary product values during the search process (e.g., preference discovery).
    - Other index-valued behaviors. Requiring Independence assumption. (e.g. product discovery).
- Key point: we focus on its impact on the ranking, but not what new optimal rules it introduces.

Extension 2: Search and Product Discovery (Greminger 2022)

- Take the search-and-product-discovery model as an example.
- The consumer has partial knowledge of the alternatives in the choice set $\mathcal{M}_i$. She can pay a discovery cost (c^{dis}) to discover more alternatives with uncertainty.
- Greminger (2022) proves that the discovery behavior has an independent and invariant discovery value (DV). Consider the value of dth discovery on route r also follows an additive form:

$$v_{ird} = \Theta_i(E_r(X^1_{ijr}), Var_r(X^1_{ijr}), c^{ins}_{ijr}, c^{dis}_{ir}) + \tau_{ird}, \quad \text{where } \Pr(\tau_{ird} < x) = F^{\tau}(x)$$

- Each step: Buy inspected (end search), inspect uninspected, or discover through one of many routes to find more uninspected products.
- Discovery changes the rank conditions by expanding $\mathcal{M}_i$.

Extension 2: PR-GHK Simulation

- KSFS is still applicable (Zhang et al., 2023) but is more challenging in practice due to the increased model complexity.
- Greminger (2024) purposed an OSR-GHK estimator that does not employ full search path information, as it is observed in his specification.
- PR-GHK idea: specify a multi-layer ranking condition of u_{ih}, z_{ij}, and v_{ir} from the data.
  - Segment the search process into sessions with discoveries. Each session has a stable linear ranking.
  - Take the DV of each session as the 'sub-core value' of each session.
  - State the ranking condition of the last session as the bottom, and lay the other conditions over up.
- Monte Carlo Simulation Results (100 reps, 2000 consumers):

	True val	PR-GHK	PR-GHK		True val	PR-GHK	PR-GHK
β_1:	2.00	2.17 (0.20)	1.94 (0.04)	$\log c_{ins}$	-2.00	-1.98 (0.04)	-2.00 (0.03)
β_2:	1.00	1.36 (0.25)	0.96 (0.05)	$\log c_{dis}$	-2.00	-1.87 (0.04)	-1.96 (0.04)
β_3:	-0.55	-0.48 (0.15)	-0.53 (0.02)	Draws		200	1000

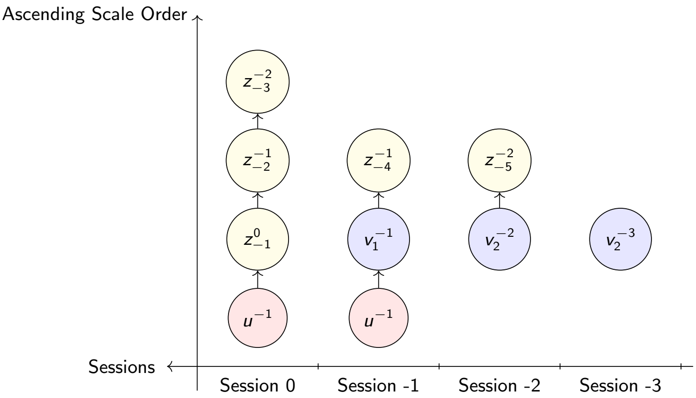
Conclusion

- This paper proposes a structure of the optimal solution to the Sequential Search Model that is more empirically friendly.
- Easy for specifying identification argument and implementing estimation without information loss.
- Very flexible for partial or additional information. Fits for a wide range of model variations with the independence assumption.
- Suitable for policy evaluations of consumers' search behavior.

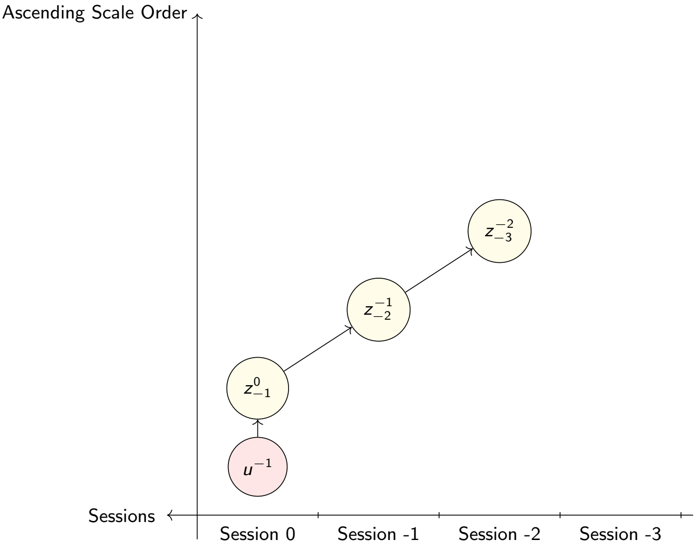
SPD Implementation Illustration

- Suppose initially 2 products are available, 2 routes, every inspection discover 2 prods.
- Consider the following sequence: $\{D2 \mid S3, D2 \mid S5, D1 \mid S4, S6, S7, P5\}$.
- Number the sessions from backward as 0, -1, -2, -3. Number the inspections from backward as -1, -2, -3, -4, -5.
- Define the values of behaviors as follows:
  - For purchasing: u^a, a is the session number in which the MV of the purchased product is realized (inspected).
  - For inspection: z_b^c, b is the inspection number, c is the session number in which the RV of the inspected product is realized (discovered).
  - For discovery: v_d^e, d is the route number, e is the session number where the DV is realized.
- Sub-core values for previous sessions: DV; for session 0: $\min\{u_a, z_b^0\}$
- Construct the ranking condition for each session sequentially.
- Core value for each session: minimum among all subsequent sub-core values.

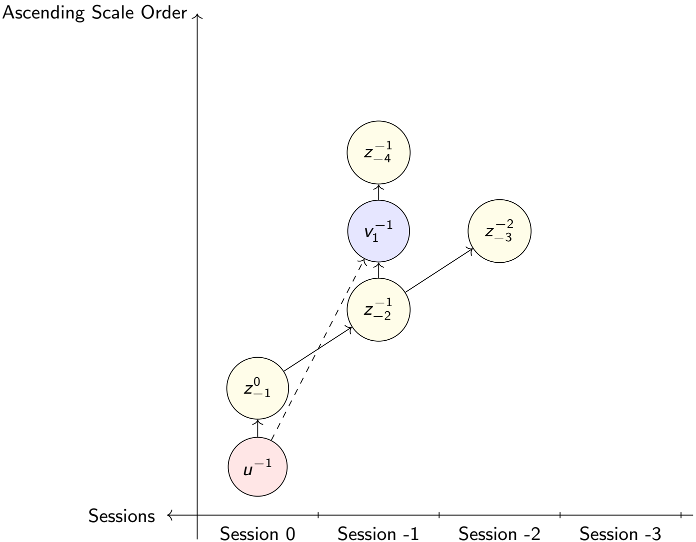
SPD Implementation Illustration



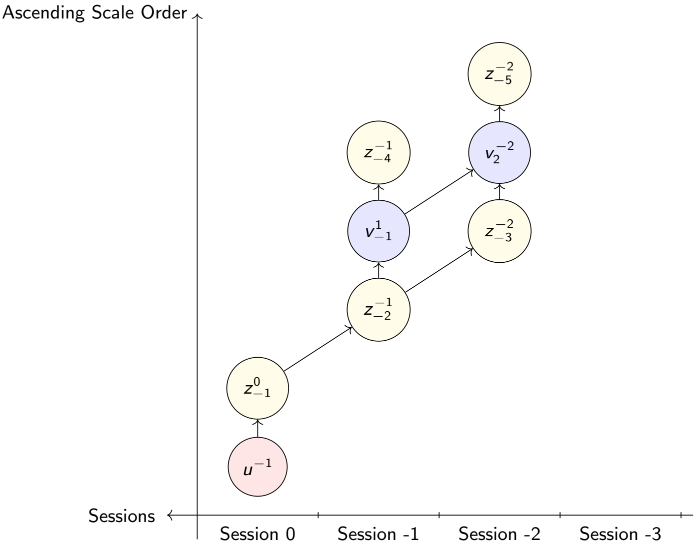
SPD Implementation Illustration



SPD Implementation Illustration



SPD Implementation Illustration



SPD Implementation Illustration