Problem Set: Stochastic Knapsack and Matching

Viswanath Nagarajan

1. **All-or-Nothing Stochastic Knapsack.** Consider the basic stochastic knapsack setting with n jobs, where each job $i \in [n]$ has deterministic reward r_i and a random size S_i (of known distribution π_i). We are also given a budget B. A subset $S \subseteq [n]$ of jobs must be chosen up-front. Then these jobs are executed, and there are two possibilities for the total reward accrued:

   - (*All*) If the total size of jobs in S is at most B then we obtain reward $\sum_{i \in S} r_i$.
   - (*Nothing*) If the total size of jobs in S is greater than B then we obtain zero reward.

   Provide a constant-factor approximation algorithm.

2. **Ordered Stochastic Knapsack.** Consider again the basic stochastic knapsack setting: there are n jobs, each job $i \in [n]$ has deterministic reward r_i and a random size S_i (of known distribution π_i), and there is a deadline B. All this information is available up-front. The n jobs will arrive in an *adversarial* order, and we want to design an "admission control" algorithm that accepts or rejects each job upon its arrival. The reward obtained is the total reward of all jobs that complete before the deadline B.

   - Give an algorithm (for this ordered setting) achieving expected reward $\Omega(1)$ times the optimal expected reward of the stochastic knapsack instance where jobs are not restricted to be considered in the given order.

     This shows that the "order" restriction in stochastic knapsack does not affect the optimal expected reward by more than a constant factor.

   - Now consider the setting where the job arrival order is known up-front. Assume also that all sizes are integers between 0 and $B + 1$. Provide a polynomial (in n, B) time adaptive algorithm.

3. **Stochastic Knapsack with Cancellation.** Consider again the stochastic knapsack setting. Now we have the option of canceling long-running jobs: if a job is canceled it yields no reward and it can not be executed again.

   - Provide an instance where the optimal expected reward with cancellation is much larger than the optimum without cancellation.
   - Give an $O(1)$-approximation algorithm for stochastic knapsack with cancellation.
   - Does your algorithm extend to the *correlated* stochastic knapsack with cancellation?

4. **Stochastic Multiple Knapsack.** We are given a set of n jobs. Each job $i \in [n]$ has deterministic reward r_i and a random ℓ-dimensional size $S_i \in \mathbb{R}_+^\ell$ (this may denote the usage of different resources, eg. time, fuel). The sizes of different jobs are independent, and we

have access to all the probability distributions. We also have an ℓ-dimensional budget vector $B \in \mathbb{R}_+^\ell$. Provide an approximation algorithm for this multi-dimensional stochastic knapsack problem. Your approximation ratio may depend on ℓ.

5. **Stochastic Orienteering with Distinct Budgets.** We are given a set of n jobs located at distinct vertices of a finite metric (V, d). Each job $v \in V$ has deterministic reward r_v and a random processing-time S_v (as in stochastic knapsack). The goal is to compute a path (possibly adaptive) originating from a depot $w \in V$ that visits vertices and runs the respective jobs. There are two budget constraints: a limit B on the total processing time, and a limit D on the total distance traveled. Note that the distances are deterministic.

   Assume an α-approximation algorithm for the *deterministic orienteering* problem: given metric (V, d) with reward p_v at each vertex $v \in V$, depot w and budget D, find a path of length at most D originating from w that maximizes the total reward.

   Give an $O(\alpha)$-approximation algorithm for this two-budget stochastic orienteering problem.

6. **Safe and Unsafe Stochastic Matching Policies.** Consider the stochastic matching setting: graph $G = (V, E)$ where each edge is active/present independently with probability p_e, and where each vertex $v \in V$ has a "patience bound" t_v. If an edge is probed and found active then it must be included in the current matching. The goal here is to probe a sequence of edges (satisfying the patience/matching constraints) so as to maximize the expected *number of probed edges*. Note the difference from the usual matching-size objective.

   A *safe* policy is one where an edge may be probed only if there is *zero* probability of violating the patience or matching constraints. These are the kind of policies we saw in the lecture.

   An *unsafe* policy is one where an edge may be probed even if there is positive probability of violating some constraint; however if a matching/patience constraint is violated then the policy ends, and the offending edge does not contribute to the objective.

   How do optimal safe and unsafe policies compare? I.e. provide upper and lower bounds for:

$$\frac{OPT_{unsafe}}{OPT_{safe}}.$$

7. **Stochastic Matching in Rounds.** Consider the stochastic matching setting as above: graph $G = (V, E)$ where each edge is active independently with probability p_e, and where each vertex $v \in V$ has a "patience bound" t_v. Suppose that edges may now be probed in *rounds*: in each round any *matching of size at most* k can be probed simultaneously. There are a total of b rounds. Find an approximation algorithm for maximizing the expected number of matched edges.

8. **Stochastic Matching with Deadlines.** Consider a set of n donor-patients in a kidney-exchange. For each pair $u, v \in [n]$ there is a probability $p_{u,v}$ of being compatible. On each day, the exchange can perform one compatibility test: if the pair is found to be compatible they get matched immediately and leave the exchange. Each donor-patient $i \in [n]$ will be in the system for d_i days. The goal is to schedule tests (possibly adaptively) so as to maximize the expected number of matched pairs. Provide an $O(1)$-approximation algorithm.