An Algorithm for Approximating Piecewise Linear Concave Functions from Sample Gradients

Huseyin Topaloglu, Warren Powell*

July 31, 2002

Abstract

An effective algorithm for solving stochastic resource allocation problems is to build piecewise linear, concave approximations of the recourse function based on sample gradient information. Algorithms based on this approach are proving useful in application areas such as the newsvendor problem, physical distribution and fleet management. These algorithms require the adaptive estimation of the approximations of the recourse function that maintain concavity at every iteration. In this paper, we prove convergence for a particular version of an algorithm that produces approximations from stochastic gradient information while maintaining concavity.

Keywords: Stochastic gradient methods, approximation, almost sure convergence, concave functions

Consider the following optimization problem:

$$\max_{x \in \mathcal{X}} \quad F(x) = E\left[f(x, \xi)\right]$$

where ξ is a random variable defined on the probability space $(\Omega, \mathcal{H}, \mathcal{P})$. We assume that $\mathcal{X} \subset \mathbb{R}_+$, $f(\cdot, \xi)$ is linear on the intervals $[s-1, s]$ for $s \in \{1, \ldots, S\}$ and concave for every realization of ξ. Problems of this kind arise in two-stage stochastic programs with separable recourse.

In these problem classes computing the function $F(\cdot)$ is usually intractable. However it is relatively easy to compute the function $f(\cdot, \xi)$ for a certain realization of ξ or compute the slope of $f(\cdot, \xi)$ at the point s given by $f(s, \xi) - f(s-1, \xi)$. In this paper we present a sampling-based method that constructs approximations of the function $F(\cdot)$ using slopes of the function $f(\cdot, \xi)$ at different points in the domain and for different realizations of ξ. We intend to solve sequences of approximations, and as a result it is important that we retain the concavity property at every iteration.

*Department of Operations Research and Financial Engineering, Princeton University, Princeton, NJ 08544 powell@princeton.edu

One application of this approach is the CAVE algorithm for the newsvendor problem (Godfrey & Powell (2001)) that constructs piecewise linear approximations using only sample gradient information. Experimental work shows that it provides very high quality solutions, without requiring that the demands follow a particular distribution, and using no information other than whether the inventory was depleted at the end of the day. Related research suggests that nonlinear functional approximations based on sample gradient information can produce useful results for general stochastic resource allocation problems (see, for example, Godfrey & Powell (to appear)).

In the stochastic programming literature this idea emerges in the context of stochastic gradient methods. These approaches generate a sequence of solutions that converge to an optimal solution. Some of them only generate the sequence without providing an approximation of the recourse function (Gupal & Bazhenov (1972), Ermoliev (1988), Birge & Qi (1995) and Birge, Qi & Wei (1998)), whereas some others attempt to approximate the recourse function explicitly (Higle & Sen (1991), Au, Higle & Sen (1994), Chen & Powell (1999), Godfrey & Powell (2001)).

In this paper, we provide a proof of convergence of a variant of the CAVE algorithm used for the discrete newsvendor problem, under the assumption that we sample all slopes with positive probability. If we did not maintain concavity at each iteration, this result would be immediate. This paper proves convergence in the presence of steps used to maintain concavity.

In section 1, we define the basic notation and present the algorithm for constructing the approximations. In section 2, we present our proof technique and establish the convergence of our approximation scheme.

1 Basic Notation

We have a probability space $(\Omega, \mathcal{H}, \mathcal{P})$ and on this space we have a generic random variable ξ, possibly with $\xi(\omega) = \omega$. We assume $f(\cdot, \xi)$ is linear on the intervals $[s-1, s]$ for all $s \in \{1, \ldots, S\}$ and concave. We let $F(s) = E\left[f(s, \xi)\right]$ for all $s \in \{0, \ldots, S\}$.

We adopt an alternative way of describing $F(\cdot)$ by using the increments of the function on the intervals $\{[s-1, s] : s \in \{1, \ldots, S\}\}$. We define the following:

$$v_s(\omega) = f(s, \omega) \text{-} f(s-1, \omega),\ s \in \{1, 2, \ldots, S\},\ \omega \in \Omega.$$

$$v_s = E\left[v_s(\xi)\right].$$

Clearly, for $s \in \{0, 1, 2, \ldots, S\}$, $f(s, \omega)$ can be written as:

$$f(s, \omega) = f(0, \omega) + v_1(\omega) + v_2(\omega) + \ldots + v_s(\omega).$$

By taking the expectation of both sides, we get:

$$F(s) = F(0) + v_1 + v_2 + \ldots + v_s.$$

Therefore, instead of approximating the values of $F(s)$ for $s \in \{0, \ldots, S\}$, we approximate the values of v_s for $s \in \{1, \ldots, S\}$ (From an optimization point of view, the approximation of the constant $F(0)$ bears no significance).

The concavity of $f(\cdot, \omega)$ and $F(\cdot)$ implies

$$\begin{aligned} v_1(\omega) \geq v_2(\omega) \geq \ldots \geq v_S(\omega) & \qquad \text{for all } \omega \in \Omega \\ v_1 \geq v_2 \geq \ldots \geq v_S. & \end{aligned}$$

We approximate the function $F(s) = F(0) + \sum_{i:1}^{s} v_i$, by $F(0) + \sum_{i:1}^{s} \hat{v}_i$ and we want our approximations to follow the concavity property of the function $F(\cdot)$. Therefore for any approximation characterized by $\{\hat{v}_1, \ldots, \hat{v}_S\}$, we need to have $\hat{v}_1 \geq \hat{v}_2 \geq \ldots \geq \hat{v}_S$.

For our problems, $v_s(\omega)$ is easy to calculate, while its expectation is generally quite difficult. We might try to estimate this expectation through updates of the form:

$$\hat{v}_s^{k+1} \quad = \quad \alpha^k \hat{v}_s^k(\omega) + (1 - \alpha^k) v_s^k$$

where $\alpha^k \in [0, 1]$ is the stepsize and $\hat{v}_s^k(\omega)$ is a sample realization of $f(s, \cdot) - f(s-1, \cdot)$ at iteration k. Under the standard assumptions $\sum_k \alpha^k = \infty$ and $\sum_k (\alpha^k)^2 < \infty$, we can easily prove that $\hat{v}_s^k$ converges to v_s *a.s.*

We describe the algorithm to construct and iteratively update the approximations in figure 1. At iteration k, $F(0) + \sum_{i:1}^{s} \hat{v}_i^k$ gives an approximation of $F(s)$.

At this point it is useful to give a numerical example. Assume $S = 5$ and the approximation at the current iteration k is characterized by:

$$\left[\begin{array}{ccccc} \hat{v}_1^k & \hat{v}_2^k & \hat{v}_3^k & \hat{v}_4^k & \hat{v}_5^k \end{array} \right] \quad = \quad \left[\begin{array}{ccccc} 5 & 3 & 2 & 0 & -1 \end{array} \right].$$

Assume $\alpha^k = 0.5$, $s^k(\omega) = 3$, $v^k(\omega) = 6$. Then,

$$\hat{v}_3^{k+1} \quad = \quad 0.5 \times 6 + (1 - 0.5) \times 2 = 4,$$

Step 0. Initialization: Choose initial values for $\hat{v}_s^0$ for all $s \in \{1, 2, \ldots, S\}$ such that $\hat{v}_1^0 \geq \hat{v}_2^0 \geq \ldots \geq \hat{v}_S^0$ that characterizes the initial approximation. Set iteration counter $k = 0$.

Step 1. Obtain sample information: Sample an outcome ω and a state $s^k(\omega) \in \{1, 2, \ldots, S\}$. Obtain the sample gradient information $v^k(\omega) = f(s^k(\omega), \omega) - f(s^k(\omega) - 1, \omega)$.

Step 2. Update the approximation: For all $j \in \{1, \ldots, S\}$, set:

$$\hat{v}_j^{k+1} = \begin{cases} \alpha^k \ v^k(\omega) + (1 - \alpha^k) \ \hat{v}_j^k & \text{if } s^k(\omega) = j \\ \alpha^k \ v^k(\omega) + (1 - \alpha^k) \ \hat{v}_i^k & \text{if } s^k(\omega) = i < j \text{ and } \alpha^k v^k(\omega) + (1 - \alpha^k)\hat{v}_i^k < \hat{v}_j^k \\ \alpha^k \ v^k(\omega) + (1 - \alpha^k) \ \hat{v}_i^k & \text{if } s^k(\omega) = i > j \text{ and } \hat{v}_j^k < \alpha^k v^k(\omega) + (1 - \alpha^k)\hat{v}_i^k \\ \hat{v}_j^k & \text{otherwise.} \end{cases}$$

Set iteration counter $k = k + 1$. Go to step 1.

Figure 1: The algorithm to update the expected recourse function approximation

$s^k(\omega) = 3 > 2$ and $3 = \hat{v}_2^k < \alpha^k v^k(\omega) + (1 - \alpha^k)\hat{v}_3^k = 4$. Therefore $\hat{v}_2^{k+1} = \alpha^k v^k(\omega) + (1 - \alpha^k)\hat{v}_3^k = 4$. Then the approximation at iteration $k + 1$ is:

$$\begin{bmatrix} \hat{v}_1^{k+1} & \hat{v}_2^{k+1} & \hat{v}_3^{k+1} & \hat{v}_4^{k+1} & \hat{v}_5^{k+1} \end{bmatrix} = \begin{bmatrix} 5 & 4 & 4 & 0 & -1 \end{bmatrix}.$$

2 Convergence for One Dimensional Recourse Functions

On the probability space $(\Omega, \mathcal{H}, \mathcal{P})$, we define the sequences of random variables $\{s^k\}_k$ and $\{v^k\}_k$ with $s^k : \Omega \to \{1, 2, \ldots, S\}$ and $v^k : \Omega \to \mathbb{R}$. $\{\alpha^k\}_k$ is a sequence of step sizes with $0 \leq \alpha^k \leq 1$. We assume the following:

A.1. v^k is uniformly bounded. That is $|v^k| < M < \infty$.

A.2. Given s^k, the conditional distribution of v^k is the same for all k.

A.3. Let $E\left[v^k \mid s^k = i\right] = v_i$, which does not depend on k by $A.2$ with $v_1 \geq v_2 \geq \ldots \geq v_S$.

A.4. Given s^k, v^k is independent of s^n and v^n for $n < k$.

A.5. For all $j \in \{1, \ldots, S\}$, the step size sequence $\{\alpha^k\}_k$ satisfies

$$\sum_k (\alpha^k)^2 < \infty \text{ and } \sum_k 1_{\{s^k = j\}} \alpha^k = \infty \ a.s.$$

We define the sequences of random variables $\{\hat{v}_j^k\}_k$ for all $j \in \{1, 2, \ldots, S\}$ recursively as follows:

$$\hat{v}_j^{k+1}(\omega) = \begin{cases} \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_j^k(\omega) & \text{if } s^k(\omega) = j \\ \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_i^k(\omega) & \text{if } s^k(\omega) = i < j \\ & \quad \text{and } \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_i^k(\omega) < \hat{v}_j^k(\omega) \\ \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_i^k(\omega) & \text{if } s^k(\omega) = i > j \\ & \quad \text{and } \hat{v}_j^k(\omega) < \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_i^k(\omega) \\ \hat{v}_j^k(\omega) & \text{otherwise} \end{cases}$$

where $\hat{v}_j^0$ for $j \in \{1, \ldots, S\}$ are deterministic numbers satisfying $\hat{v}_1^0 \geq \hat{v}_2^0 \geq \ldots \geq \hat{v}_S^0$. We call the process of calculating the values of $\{\hat{v}_j^{k+1}(\omega) : j \in \{1, \ldots, S\}\}$ using the values of $\{\hat{v}_j^k(\omega) : j \in \{1, \ldots, S\}\}$ as an *updating process*.

We like to draw the parallel between the algorithm in figure 1 and the sequences of random variables $\{s^k\}_k$, $\{v^k\}_k$ and $\{v_j^k\}_k$. The random variable s^k corresponds to the state picked from the set $\{1, 2, \ldots, S\}$ and v^k corresponds to the gradient information obtained at iteration k in step 1. Note that by $A.3$, given that state i is chosen, the expected value of v^k is v_i which is in agreement with the notation we developed in the previous section. The random variables $\hat{v}_j^k$ for $j \in \{1, \ldots, S\}$ are the approximations of v_j at iteration k.

2.1 The Proof Technique

We are ultimately interested in showing that $\hat{v}_j^k \to v_j$ $a.s.$ for all $j \in \{1, 2, \ldots, S\}$. We use induction to show the final result. Our induction technique is motivated by the observation that instead of constructing approximations of $F(\cdot)$ over the whole domain $\{1, \ldots, S\}$ (these approximations are characterized by $\{\hat{v}_j^k\}_k$ for $j \in \{1, \ldots, S\}$), we can construct approximations on a restricted domain $\{1, \ldots, n\}$ where $n \leq S$. These restricted approximations are built simply by omitting the updating process when $s^k(\omega) > n$. We let $\{\hat{v}_j^{n,k}\}_k$ for $j \in \{1, \ldots, n\}$ be the slopes characterizing the approximations of $F(\cdot)$ restricted to the domain $\{1, \ldots, n\}$. We can formally write the updating process of the random variables $\{\hat{v}_j^{n,k}\}_k$ as follows:

$$\hat{v}_j^{n,k+1}(\omega) = \begin{cases} \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_j^{n,k}(\omega) & \text{if } s^k(\omega) = j \text{ and } s^k(\omega) \leq n \\ \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_i^{n,k}(\omega) & \text{if } s^k(\omega) = i < j \leq n \text{ and} \\ & \quad \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_i^{n,k}(\omega) < \hat{v}_j^{n,k}(\omega) \\ \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_i^{n,k}(\omega) & \text{if } n \geq s^k(\omega) = i > j \text{ and} \\ & \quad \hat{v}_j^{n,k}(\omega) < \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_i^{n,k}(\omega) \\ \hat{v}_j^{n,k}(\omega) & \text{otherwise.} \end{cases} \tag{1}$$

Note that in the updating process above, if $s^k(\omega) > n$, $\hat{v}_j^{n,k+1}(\omega) = \hat{v}_j^{n,k}(\omega)$ for all $j \in \{1, \ldots, n\}$.

We assume that $\hat{v}_j^{n,0}$ for $j \in \{1,\ldots,n\}$ are deterministic numbers satisfying

$$\hat{v}_1^{n,0} \geq \hat{v}_2^{n,0} \geq \ldots \geq \hat{v}_n^{n,0} \qquad \text{for all } n \in \{1,\ldots,S\} \tag{2}$$

$$\hat{v}_j^{n-1,0} \leq \hat{v}_j^{n,0} \qquad \text{for all } j \in \{1,\ldots,n-1\}, n \in \{2,\ldots,S\}. \tag{3}$$

Then the induction argument is based on the following two steps:

1. Show that $\hat{v}_j^{2,k} \to v_j$ $a.s.$ for all $j \in \{1,2\}$. This is the initial condition for the induction hypothesis. (Lemmas 2.4 and 2.5)

2. Assuming that $\hat{v}_j^{n,k} \to v_j$ $a.s.$ for all $j \in \{1,\ldots,n\}$, show that $\hat{v}_j^{n+1,k} \to v_j$ $a.s.$ for all $j \in \{1,\ldots,n+1\}$. (Lemma 2.6 and proposition 2.1)

In effect this proof technique assumes that the algorithm produces $a.s.$ convergent approximations of $F(\cdot)$ on the restricted domain $\{1,\ldots,n\}$ and shows that the algorithm produces $a.s.$ convergent approximations of $F(\cdot)$ on the restricted domain $\{1,\ldots,n+1\}$. In the remainder of this section, we define some events and random variables that we make use of throughout the paper.

The crucial part of the algorithm is enforcing the concavity on the approximations. In order to characterize the iterations that this occurs we define the sequences of events $\{U_{i,j}^{n,k}\}_k$ for $i,j \in \{1,\ldots,n\}$ and $n \in \{1,\ldots,S\}$ as:

$$U_{i,j}^{n,k} = \begin{cases} \left\{\omega \in \Omega : s^k(\omega) = i, \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_i^{n,k}(\omega) < \hat{v}_j^{n,k}(\omega)\right\} & \text{if } i < j \\ \left\{\omega \in \Omega : s^k(\omega) = i, \hat{v}_j^{n,k}(\omega) < \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_i^{n,k}(\omega)\right\} & \text{if } j < i \\ \emptyset & \text{if } i = j \end{cases} \tag{4}$$

$$= \begin{cases} \left\{\omega \in \Omega : s^k(\omega) = i, \hat{v}_i^{n,k+1}(\omega) < \hat{v}_j^{n,k}(\omega)\right\} & \text{if } i < j \\ \left\{\omega \in \Omega : s^k(\omega) = i, \hat{v}_j^{n,k}(\omega) < \hat{v}_i^{n,k+1}(\omega)\right\} & \text{if } j < i \\ \emptyset & \text{if } i = j. \end{cases}$$

Roughly speaking, the event $U_{i,j}^{n,k}$ is the set of outcomes ω such that at iteration k we sample the state $s^k(\omega) = i \neq j$ and we have to enforce the concavity of the approximation (i.e. the approximation restricted to $\{1,\ldots,n\}$) over the interval $[j-1,j]$.

We give a numerical example to clarify the updating process of the random variables $\{\hat{v}_j^{n,k}(\omega) : n \in \{1,\ldots,S\}, j \in \{1,\ldots,n\}\}$. Assume $S = 5$, $n = 4$, $s^k(\omega) = 3$, $v^k(\omega) = -4$, $\alpha^k = 0.5$ and

$$\begin{bmatrix} \hat{v}_1^{4,k}(\omega) & \hat{v}_2^{4,k}(\omega) & \hat{v}_3^{4,k}(\omega) & \hat{v}_4^{4,k}(\omega) & \cdot \end{bmatrix} = \begin{bmatrix} 4.5 & 2.5 & 1.5 & 0 & \cdot \end{bmatrix}.$$

Then $\omega \in U_{3,4}^{4,k}$ and

$$\begin{bmatrix} \hat{v}_1^{4,k+1}(\omega) & \hat{v}_2^{4,k+1}(\omega) & \hat{v}_3^{4,k+1}(\omega) & \hat{v}_4^{4,k+1}(\omega) & \cdot \end{bmatrix} = \begin{bmatrix} 4.5 & 2.5 & -1.25 & -1.25 & \cdot \end{bmatrix}.$$

Using the events $U_{i,j}^{n,k}$ for $i, j \in \{1, \ldots, n\}$, the updating process of the random variable $\hat{v}_j^{n,k}$ can be written as:

$$\hat{v}_j^{n,k+1}(\omega) = 1_{\{s^k(\omega)=j\}} \left[\alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_j^{n,k}(\omega)\right] + \sum_{i=1}^{n} 1_{\{\omega \in U_{i,j}^{n,k}\}} \hat{v}_i^{n,k+1}(\omega) + (1 - 1_{\{s^k(\omega)=j\}} - \sum_{i=1}^{n} 1_{\{\omega \in U_{i,j}^{n,k}\}}) \hat{v}_j^{n,k}(\omega) \tag{5}$$

where we use 1_H to denote the indicator random variable

$$1_H(\omega) = \begin{cases} 1 & \text{if } \omega \in H \\ 0 & \text{otherwise.} \end{cases}$$

The logic behind the equation above is fairly simple. When computing the value of the random variable $\hat{v}_j^{n,k+1}(\omega)$ by using the values of $\{\hat{v}_j^{n,k}(\omega) : j \in \{1, \ldots, n\}\}$ there are three cases to consider: The first one is the case where $s^k(\omega) = j$, the second one is the case where $s^k(\omega) = i \neq j$ but $\hat{v}_j^{n,k+1}(\omega) = \hat{v}_i^{n,k+1}(\omega)$ (i.e. $\omega \in U_{i,j}^{n,k}$) and the third one is the case where neither of these conditions hold. The three terms in the equation above reflect these three cases respectively.

Finally we define the following sequences of random variables $\{y^{n,k}\}_k$ for $n \in \{1, \ldots, S\}$:

$$y^{n,k+1}(\omega) = \begin{cases} \alpha^k v^k(\omega) + (1-\alpha^k) y^{n,k}(\omega) & \text{if } s^k(\omega) = n \\ y^{n,k}(\omega) & \text{otherwise.} \end{cases} \tag{6}$$

The sequence $\{y^{n,k}\}_k$ has a convergence property (lemma 2.1) and provides bounds on the sequence $\{\hat{v}_n^{n,k}\}_k$ (lemma 2.3). These bounds are instrumental in the proofs of lemmas 2.4 and 2.6. We assume that $y^{n,0}$ for $n \in \{1, \ldots, S\}$ are deterministic numbers satisfying

$$y^{n,0} \geq \hat{v}_n^{n,0} \text{ for all } n \in \{1, \ldots, S\}. \tag{7}$$

2.2 Important Properties

In this section we show useful properties of the random variables we defined. The first property is the fact that the restricted approximations are also concave.

Property 2.1 *For any k,*

$$\hat{v}_1^{n,k}(\omega) \geq \hat{v}_2^{n,k}(\omega) \geq \ldots \geq \hat{v}_{n-1}^{n,k}(\omega) \geq \hat{v}_n^{n,k}(\omega)$$

for all $n \in \{1, \ldots, S\}$ and $\omega \in \Omega$.

The second property we show is the fact that the approximations restricted to the domain $\{1, \ldots, S\}$ is the same as the unrestricted approximations.

Property 2.2 $\hat{v}_j^{S,k}(\omega) = \hat{v}_j^k(\omega)$ *for all* $j \in \{1, \ldots, S\}$ *and* $\omega \in \Omega$.

In lemmas 2.1, 2.2 and 2.3 we show some useful properties of the sequences of random variables $\{\hat{v}_j^{n,k}\}_k$ and $\{y^{n,k}\}_k$.

Lemma 2.1 *Assume that assumptions A.1-A.5 hold.* $y^{n,k} \to v_n$ *a.s. for all* $n \in \{1, \ldots S\}$.

Proof: The proof is obvious by noting the updating procedure for $y^{n,k}(\omega)$ and the assumption that $\sum_k (\alpha^k)^2 < \infty$ and $\sum_k 1_{\{s^k=j\}}\alpha^k = \infty$ *a.s.* $\square$

Lemma 2.2 *Assume that assumptions A.1-A.5 hold. Then* $\hat{v}_1^{1,k} \to v_1$ *a.s.*

Proof: The proof is similar to the proof of lemma 2.1 by noting that

$$\hat{v}_1^{1,k+1}(\omega) = \begin{cases} \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_1^{1,k}(\omega) & \text{if } s^k(\omega) = 1 \\ \hat{v}_1^{1,k}(\omega) & \text{otherwise.} \end{cases} \quad \square$$

Lemma 2.3 *Fix* $n \in \{2, \ldots, S\}$. *Then for all* $\omega \in \Omega$ *and* k, *we have*

$$\begin{aligned} \hat{v}_j^{n-1,k}(\omega) &\leq \hat{v}_j^{n,k}(\omega) \text{ for all } j \in \{1, \ldots, n-1\} \\ y^{n,k}(\omega) &\geq \hat{v}_n^{n,k}(\omega). \end{aligned}$$

Proof: We only show the first inequality by using induction. The proof requires a case by case analysis. We only show one case and state the remaining cases to be considered. For $k = 0$, the result holds by (3). Fix $\omega \in \Omega$ and assume for arbitrary k we have $\hat{v}_j^{n-1,k}(\omega) \leq \hat{v}_j^{n,k}(\omega)$ for all $j \in \{1, \ldots, n-1\}$. We fix $j \in \{1, \ldots, n-1\}$ and consider five cases:

[1] If $1 \leq s^k(\omega) < j$: There are four subcases to consider since $1 \leq s^k(\omega) < j < n$:

[1.a] $\hat{v}_{s^k}^{n,k+1}(\omega) < \hat{v}_j^{n,k}(\omega)$ and $\hat{v}_{s^k}^{n-1,k+1}(\omega) < \hat{v}_j^{n-1,k}(\omega)$: By the updating procedure we have $\hat{v}_j^{n,k+1}(\omega) = \hat{v}_{s^k}^{n,k+1}(\omega)$ and $\hat{v}_j^{n-1,k+1}(\omega) = \hat{v}_{s^k}^{n-1,k+1}(\omega)$. By the induction hypothesis we have $\hat{v}_{s^k}^{n-1,k}(\omega) \leq \hat{v}_{s^k}^{n,k}(\omega)$. Then $\hat{v}_{s^k}^{n-1,k+1}(\omega) = \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_{s^k}^{n-1,k}(\omega) \leq \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_{s^k}^{n,k}(\omega) = \hat{v}_{s^k}^{n,k+1}(\omega)$. Putting these three results together, we obtain $\hat{v}_j^{n-1,k+1}(\omega) = \hat{v}_{s^k}^{n-1,k+1}(\omega) \leq \hat{v}_{s^k}^{n,k+1}(\omega) = \hat{v}_j^{n,k+1}(\omega)$.

[1.b] $\hat{v}_{s^k}^{n,k+1}(\omega) < \hat{v}_j^{n,k}(\omega)$ and $\hat{v}_{s^k}^{n-1,k+1}(\omega) \geq \hat{v}_j^{n-1,k}(\omega)$.

[1.c] $\hat{v}_{s^k}^{n,k+1}(\omega) \geq \hat{v}_j^{n,k}(\omega)$ and $\hat{v}_{s^k}^{n-1,k+1}(\omega) < \hat{v}_j^{n-1,k}(\omega)$.

[1.d] $\hat{v}_{s^k}^{n,k+1}(\omega) \geq \hat{v}_j^{n,k}(\omega)$ and $\hat{v}_{s^k}^{n-1,k+1}(\omega) \geq \hat{v}_j^{n-1,k}(\omega)$.

The remaining cases to consider are:

[2] $s^k(\omega) = j$, [3] $j < s^k(\omega) \leq n-1$, [4] $s^k(\omega) = n$, and [5] $s^k(\omega) \geq n+1$. $\square$

2.3 Convergence Proof

In order to make the core idea of the proof clear we momentarily assume that $v_1 > v_2 > \ldots > v_S$ instead of $v_1 \geq v_2 \geq \ldots \geq v_S$. We show how to drop this assumption later. In lemmas 2.4 and 2.5, we concentrate on the case when $n = 2$ to establish the initial conditions for our induction argument.

Lemma 2.4 *Assume that assumptions A.1-A.5 hold. Then* $\sum_k 1_{U_{1,2}^{2,k}} < \infty$ *and* $\sum_k 1_{U_{2,1}^{2,k}} < \infty$ *a.s.*

Proof: Using (4), we can write:

$$U_{1,2}^{2,k} = \left\{\omega \in \Omega : s^k(\omega) = 1, \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_1^{2,k}(\omega) < \hat{v}_2^{2,k}(\omega)\right\}.$$

In the following, the second and third lines are implied by lemma 2.3 and the fourth line is implied by the updating process of $\hat{v}_1^{1,k}$:

$$\begin{aligned}
U_{1,2}^{2,k} &= \left\{\omega \in \Omega : s^k(\omega) = 1, \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_1^{2,k}(\omega) < \hat{v}_2^{2,k}(\omega)\right\} \\
&\subset \left\{\omega \in \Omega : s^k(\omega) = 1, \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_1^{1,k}(\omega) < \hat{v}_2^{2,k}(\omega)\right\} \\
&\subset \left\{\omega \in \Omega : s^k(\omega) = 1, \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_1^{1,k}(\omega) < y^{2,k}(\omega)\right\} \\
&= \left\{\omega \in \Omega : s^k(\omega) = 1, \hat{v}_1^{1,k+1}(\omega) < y^{2,k}(\omega)\right\} \\
&\subset \left\{\omega \in \Omega : \hat{v}_1^{1,k+1}(\omega) < y^{2,k}(\omega)\right\}.
\end{aligned}$$

Thus $\sum_k 1_{\{\omega \in U_{1,2}^{2,k}\}} \leq \sum_k 1_{\{\hat{v}_1^{1,k+1}(\omega) < y^{2,k}(\omega)\}}$ for all ω. From lemma 2.1 and 2.2, we have $lim_{k\to\infty}\hat{v}_1^{1,k+1}(\omega) - y^{2,k}(\omega) = v_1 - v_2$ for *a.e.* ω. Then for any $\epsilon > 0$, $\sum_k 1_{(\epsilon,\infty)} \circ |(\hat{v}_1^{1,k+1}(\omega) - y^{2,k}(\omega)) - (v_1 - v_2)| < \infty$ for *a.e.* ω. Pick $\epsilon = v_1 - v_2 > 0$:

$$\begin{aligned}
\sum_k 1_{\{\hat{v}_1^{1,k+1} < y^{2,k}\}} &= \sum_k 1_{\{(y^{2,k} - \hat{v}_1^{1,k+1}) - (v_2 - v_1) > v_1 - v_2\}} \leq \sum_k 1_{\{|(y^{2,k} - \hat{v}_1^{1,k+1}) - (v_2 - v_1)| > v_1 - v_2\}} \\
&= \sum_k 1_{(v_1 - v_2, \infty)} \circ |(\hat{v}_1^{1,k+1} - y^{2,k}) - (v_1 - v_2)| < \infty \quad a.s.
\end{aligned}$$

So $\sum_k 1_{U_{1,2}^{2,k}} \leq \sum_k 1_{\{\hat{v}_1^{1,k+1} < y^{2,k}\}} < \infty$ *a.s.*

The proof for $\sum_k 1_{U_{2,1}^{2,k}} < \infty$ *a.s.* follows the same lines. $\square$

The following lemma establishes the initial conditions for our induction argument.

Lemma 2.5 *Assume that assumptions A.1-A5 hold. Then $\hat{v}_j^{2,k} \to v_j$ a.s. for all $j \in \{1,2\}$.*

Proof: We start with showing that $\hat{v}_2^{2,k} \to v_2$ *a.s.* $y^{2,k+1}(\omega)$ can be written as:

$$y^{2,k+1}(\omega) = 1_{\{s^k(\omega)=2\}} \left[\alpha^k v^k(\omega) + (1-\alpha^k) y^{2,k}(\omega)\right] + (1 - 1_{\{s^k(\omega)=2\}})\ y^{2,k}(\omega).$$

On the other hand using (5), $\hat{v}_2^{2,k+1}(\omega)$ can be written as:

$$\begin{aligned}\hat{v}_2^{2,k+1}(\omega) &= 1_{\{s^k(\omega)=2\}} \left[\alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_2^{2,k}(\omega)\right] + 1_{\{\omega \in U_{1,2}^{2,k}\}} \hat{v}_1^{2,k+1}(\omega) \\ &\quad + (1 - 1_{\{s^k(\omega)=2\}} - 1_{\{\omega \in U_{1,2}^{2,k}\}}) \hat{v}_2^{2,k}(\omega).\end{aligned}$$

If we subtract the two equations above side by side, we get:

$$\begin{aligned}y^{2,k+1}(\omega) - \hat{v}_2^{2,k+1}(\omega) &= 1_{\{s^k(\omega)=2\}}(1-\alpha^k) \left[y^{2,k}(\omega) - \hat{v}_2^{2,k}(\omega)\right] \\ &\quad + (1 - 1_{\{s^k(\omega)=2\}}) \left[y^{2,k}(\omega) - \hat{v}_2^{2,k}(\omega)\right] - 1_{\{\omega \in U_{1,2}^{2,k}\}} \left[\hat{v}_1^{2,k+1}(\omega) - \hat{v}_2^{2,k}(\omega)\right].\end{aligned}$$

By lemma 2.4, for *a.e.* ω, there exists a finite $N(\omega)$ such that $1_{\{\omega \in U_{1,2}^{2,k}\}} = 0$ for all $k \geq N(\omega)$. Then for $k \geq N(\omega)$:

$$\left|y^{2,k+1}(\omega) - \hat{v}_2^{2,k+1}(\omega)\right| = \left[1_{\{s^k(\omega)=2\}}(1-\alpha^k) + (1 - 1_{\{s^k(\omega)=2\}})\right] \left|y^{2,k}(\omega) - \hat{v}_2^{2,k}(\omega)\right|$$

$$\left|y^{2,k+1}(\omega) - \hat{v}_2^{2,k+1}(\omega)\right| - \left|y^{2,k}(\omega) - \hat{v}_2^{2,k}(\omega)\right| = -1_{\{s^k(\omega)=2\}} \alpha^k \left|y^{2,k}(\omega) - \hat{v}_2^{2,k}(\omega)\right|.$$

If we write the equation above for $k = N(\omega)$ to $k = K > N(\omega)$ and add them side by side, we get:

$$\left|y^{2,K+1}(\omega) - \hat{v}_2^{2,K+1}(\omega)\right| - \left|y^{2,N(\omega)}(\omega) - \hat{v}_2^{2,N(\omega)}(\omega)\right| = -\sum_{k=N(\omega)}^{K} 1_{\{s^k(\omega)=2\}} \alpha^k \left|y^{2,k}(\omega) - \hat{v}_2^{2,k}(\omega)\right|.$$

Since $\{v^k\}_k$ are uniformly bounded, if we take the limit of the left side as $K \to \infty$, the left side is bounded. So must be the right side. But by assumption $\sum_{k=N(\omega)}^{\infty} 1_{\{s^k(\omega)=2\}} \alpha^k = \infty$ for *a.e.* ω since $N(\omega)$ is finite for *a.e.* ω. Then $\lim_{k\to\infty} \left|y^{2,k}(\omega) - \hat{v}_2^{2,k}(\omega)\right| \to 0$ for *a.e.* ω. $y_2^{2,k} \to v_2$ *a.s.* by lemma 2.1, hence $\hat{v}_2^{2,k} \to v_2$ *a.s.*

The proof for $\hat{v}_1^{2,k}$ follows the same lines by using lemma 2.2 and noting that

$$\begin{aligned}\hat{v}_1^{1,k+1}(\omega) &= 1_{\{s^k(\omega)=1\}} \left[\alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_1^{1,k}(\omega)\right] + (1 - 1_{\{s^k(\omega)=1\}})\ \hat{v}_1^{1,k}(\omega) \\ \hat{v}_1^{2,k+1}(\omega) &= 1_{\{s^k(\omega)=1\}} \left[\alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_1^{2,k}(\omega)\right] + 1_{\{\omega \in U_{2,1}^{2,k}\}} \hat{v}_2^{2,k+1}(\omega) \\ &\quad + (1 - 1_{\{s^k(\omega)=1\}} - 1_{\{\omega \in U_{2,1}^{2,k}\}}) \hat{v}_1^{2,k}(\omega)\ . \square\end{aligned}$$

Lemma 2.5 establishes the initial condition that we need for our induction argument. We now present the induction argument to complete the proof in lemma 2.6 and proposition 2.1.

Lemma 2.6 *Assume that assumptions A.1-A.5 hold. Also assume that* $\hat{v}_j^{n,k} \to v_j$ *a.s. for all* $j \in \{1,\ldots,n\}$. *Then* $\sum_k 1_{U_{j,n+1}^{n+1,k}} < \infty$ *and* $\sum_k 1_{U_{n+1,j}^{n+1,k}} < \infty$ *a.s. for all* $j \in \{1,\ldots,n+1\}$.

Proof: Proof is similar to that of lemma 2.4 by noting that

$$\begin{aligned}
U_{j,n+1}^{n+1,k} &= \left\{\omega \in \Omega : s^k(\omega) = j, \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_j^{n+1,k}(\omega) < \hat{v}_{n+1}^{n+1,k}(\omega)\right\} \\
&\subset \left\{\omega \in \Omega : s^k(\omega) = j, \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_j^{n,k}(\omega) < \hat{v}_{n+1}^{n+1,k}(\omega)\right\} \\
&\subset \left\{\omega \in \Omega : s^k(\omega) = j, \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_j^{n,k}(\omega) < y^{n+1,k}(\omega)\right\} \\
&= \left\{\omega \in \Omega : s^k(\omega) = j, \hat{v}_j^{n,k+1}(\omega) < y^{n+1,k}(\omega)\right\} \\
&\subset \left\{\omega \in \Omega : \hat{v}_j^{n,k+1}(\omega) < y^{n+1,k}(\omega)\right\}.
\end{aligned}$$

The result follows by the assumption that $\hat{v}_j^{n,k} \to v_j$ *a.s.* for all $j \in \{1,\ldots,n\}$ and lemma 2.1. $\Box$

Roughly speaking $\sum_k 1_{U_{n+1,j}^{n+1,k}} < \infty$ *a.s.* for all $j \in \{1,\ldots,n\}$ shows that for *a.e.* ω after a finite number of iterations, if we sample $s^k(\omega) = n+1$, then we never need to enforce the concavity of the approximation (more correctly the approximation restricted to $\{1,\ldots,n+1\}$) over $[j-1,j]$. Therefore whenever $s^k(\omega) = n+1$, the approximation restricted to $\{1,\ldots,n+1\}$ remains the same from iteration k to $k+1$ over $[0,n]$. The only part of the approximation that changes is over $[n,n+1]$.

Proposition 2.1 *Assume that assumptions A.1-A.5 hold. Also assume that* $\hat{v}_j^{n,k} \to v_j$ *a.s. for all* $j \in \{1,\ldots,n\}$. *Then* $\hat{v}_j^{n+1,k} \to v_j$ *a.s. for all* $j \in \{1,\ldots,n+1\}$.

Proof: Fix $j \in \{1,\ldots,n\}$. The updating process of $\hat{v}_j^{n,k}$ is given by:

$$\hat{v}_j^{n,k+1}(\omega) = \begin{cases} \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_j^{n,k}(\omega) & \text{if } s^k(\omega) = j \text{ and } s^k(\omega) \le n \\ \hat{v}_i^{n,k+1}(\omega) & \text{if } s^k(\omega) = i < j \le n \text{ and } \hat{v}_i^{n,k+1}(\omega) < \hat{v}_j^{n,k}(\omega) \\ \hat{v}_i^{n,k+1}(\omega) & \text{if } n \ge s^k(\omega) = i > j \text{ and } \hat{v}_j^{n,k}(\omega) < \hat{v}_i^{n,k+1}(\omega) \\ \hat{v}_j^{n,k}(\omega) & \text{otherwise.} \end{cases}$$

On the other hand the updating process for $\hat{v}_j^{n+1,k}$ can be written as:

$$\hat{v}_j^{n+1,k+1}(\omega) = \begin{cases} \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_j^{n+1,k}(\omega) & \text{if } s^k(\omega) = j \text{ and } s^k(\omega) \le n+1 \\ \hat{v}_i^{n+1,k+1}(\omega) & \text{if } s^k(\omega) = i < j \le n+1 \text{ and } \hat{v}_i^{n+1,k+1}(\omega) < \hat{v}_j^{n+1,k}(\omega) \\ \hat{v}_i^{n+1,k+1}(\omega) & \text{if } n+1 \ge s^k(\omega) = i > j \text{ and } \hat{v}_j^{n+1,k}(\omega) < \hat{v}_i^{n+1,k+1}(\omega) \\ \hat{v}_j^{n+1,k}(\omega) & \text{otherwise.} \end{cases}$$

Then we can make the following three modifications to the definition of the updating process of $\hat{v}_j^{n+1,k}$:

[1] Since $j \leq n$, the condition of the first case "if $s^k(\omega) = j$ and $s^k(\omega) \leq n+1$", can be replaced by "if $s^k(\omega) = j$ and $s^k(\omega) \leq n$.

[2] Since $j \leq n$, the condition of the second case "if $s^k(\omega) = i < j \leq n+1$ and $\hat{v}_i^{n+1,k+1}(\omega) < \hat{v}_j^{n+1,k}(\omega)$" can be replaced by "if $s^k(\omega) = i < j \leq n$ and $\hat{v}_i^{n+1,k+1}(\omega) < \hat{v}_j^{n+1,k}(\omega)$".

[3] By lemma 2.6, we know that for all $j \in \{1, \ldots, n\}$ and for $a.e$ ω, there exists a finite $N(\omega)$ such that $1_{\{\omega \in U_{n+1,j}^{n+1,k}\}} = 0$ for all $k \geq N(\omega)$. Then for $a.e.$ ω and $k \geq N(\omega)$, if $s^k(\omega) = n+1$, we need to have $\hat{v}_j^{n+1,k}(\omega) \geq \hat{v}_{n+1}^{n+1,k+1}(\omega)$, since $\omega \notin U_{n+1,j}^{n+1,k} = \{\omega \in \Omega : s^k(\omega) = n+1, \hat{v}_j^{n+1,k}(\omega) < \hat{v}_{n+1}^{n+1,k+1}(\omega)\}$. Therefore, the condition of the third case never holds if $s^k(\omega) = n+1$, for $k \geq N(\omega)$. Then the condition of the third case can be replaced by "if $n \geq s^k(\omega) = i > j$ and $\hat{v}_j^{n+1,k}(\omega) < \hat{v}_i^{n+1,k+1}(\omega)$".

Thus, for $a.e.$ ω and $k \geq N(\omega)$, the updating process of $\hat{v}_j^{n+1,k}(\omega)$ is given by:

$$\hat{v}_j^{n+1,k+1}(\omega) \quad = \quad \begin{cases} \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_j^{n+1,k}(\omega) & \text{if } s^k(\omega) = j \text{ and } s^k(\omega) \leq n \\ \hat{v}_i^{n+1,k+1}(\omega) & \text{if } s^k(\omega) = i < j \leq n \text{ and } \hat{v}_i^{n+1,k+1}(\omega) < \hat{v}_j^{n+1,k}(\omega) \\ \hat{v}_i^{n+1,k+1}(\omega) & \text{if } n \geq s^k(\omega) = i > j \text{ and } \hat{v}_j^{n+1,k}(\omega) < \hat{v}_i^{n+1,k+1}(\omega) \\ \hat{v}_j^{n+1,k}(\omega) & \text{otherwise} \end{cases}$$

which is exactly the updating process of $\hat{v}_j^{n,k}(\omega)$. Since $N(\omega)$ is finite for $a.e.$ ω, $lim_{k\to\infty}\hat{v}_j^{n,k}(\omega) = \lim_{k\to\infty} \hat{v}_j^{n+1,k}(\omega)$ for $a.e.$ ω. By assumption, $\hat{v}_j^{n,k} \to v_j$ $a.s.$ for all $j \in \{1, \ldots, n\}$. Thus $\hat{v}_j^{n+1,k} \to v_j$ $a.s.$ for $j \in \{1, \ldots, n\}$.

Next, we need to show that $\hat{v}_{n+1}^{n+1,k} \to v_{n+1}$ $a.s.$ This part of the proof is very similar to the proof of lemma 2.5. □

The following theorem shows the final convergence result we are after.

Theorem 2.1 *Assume that assumptions A.1-A.5 hold. Also assume that* $v_1 > v_2 > \ldots > v_S$*. Then* $\hat{v}_j^k = \hat{v}_j^{S,k} \to v_j$ *a.s. for all* $j \in \{1, \ldots, S\}$*.*

Proof: The result follows by an induction argument using proposition 2.1, the boundary condition provided by lemma 2.5 and property 2.2. □

In the remainder of this section, we relax some of the assumptions we made to handle some cases that arise in practice.

Remark: This far we assumed that $v_1 > v_2 > \ldots > v_S$. We now assume that there exists an $s \in \{1, \ldots, S-1\}$ such that $v_s = v_{s+1}$. In this case the induction argument follows the same lines as long as $n \leq s$.

For $n = s+1$, it can be shown that for $a.e.\ \omega$, there exists a finite $N(\omega)$ such that for all $k \geq N(\omega)$, $1_{\{\omega \in U^{n,k}_{i,s+1}\}} = 0$ and $1_{\{\omega \in U^{n,k}_{s+1,i}\}} = 0$ for all $i \in \{1, \ldots, s-1\}$. Then for any $j \in \{1, \ldots, s-1\}$ and $k \geq N(\omega)$ we have:

$$\begin{aligned}\hat{v}_j^{s+1,k+1}(\omega) &= 1_{\{s^k(\omega)=j\}}\left[\alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_j^{s+1,k}(\omega)\right] + \sum_{i=1}^{s} 1_{\{\omega \in U^{s+1,k}_{i,j}\}} \hat{v}_i^{s+1,k+1}(\omega) \\ &\quad + (1 - 1_{\{s^k(\omega)=s\}} - \sum_{i=1}^{s} 1_{\{\omega \in U^{s+1,k}_{i,j}\}})\hat{v}_j^{s+1,k}(\omega).\end{aligned}$$

This is the same updating scheme as that of $\{\hat{v}_j^{s,k} : j \in \{1, \ldots, s-1\}\}$. The fact that $1_{\{\omega \in U^{n,k}_{s,s+1}\}} = 0$ and $1_{\{\omega \in U^{n,k}_{s+1,s}\}} = 0$ do not occur finitely many times $a.s.$ does not create any problems since $E\left[v^k \mid s^k = s\right] = v_s = v_{s+1} = E\left[v^k \mid s^k = s+1\right]$. Therefore the extension of the proof of the $a.s.$ convergence of the algorithm to the case where $v_1 \geq v_2 \geq \ldots \geq v_S$ is straightforward.

References

Au, K., Higle, J. & Sen, S. (1994), 'Inexact subgradient methods with applications in stochastic programming', *Mathematical Programming* **63**(1), 65–82.

Birge, J. & Qi, L. (1995), 'Subdifferential convergence in stochastic systems', *SIAM Journal on Optimization* **5**(2), 436–453.

Birge, J., Qi, L. & Wei, Z. (1998), 'Convergence analysis of some methods for minimizing a nonsmooth convex function', *Journal of Optimization Theory and Applications* **97**(2), 357–383.

Chen, Z.-L. & Powell, W. B. (1999), 'A convergent cutting-plane and partial-sampling algorithm for multistage stochastic linear programs with recourse', *Journal of Optimization Theory and Applications* **102**(3), 497–524.

Ermoliev, Y. (1988), Stochastic quasigradient methods, *in* Y. Ermoliev & R. Wets, eds, '*Numerical Techniques for Stochastic Optimization*', Springer-Verlag, Berlin.

Godfrey, G. & Powell, W. B. (to appear), 'An adaptive, dynamic programming algorithm for stochastic resource allocation problems I: Single period travel times', *Transportation Science*.

Godfrey, G. A. & Powell, W. B. (2001), 'An adaptive, distribution-free approximation for the newsvendor problem with censored demands, with applications to inventory and distribution problems', *Management Science* **47**(8), 1101–1112.

Gupal, A. M. & Bazhenov, L. G. (1972), 'A stochastic method of linearization', *Cybernetics* **8**, 482–484.

Higle, J. & Sen, S. (1991), 'Stochastic decomposition: An algorithm for two stage linear programs with recourse', *Mathematics of Operations Research* **16**(3), 650–669.

A Appendices for Referees: Proofs of Lemmas and Propositions

Proof of Lemma 2.1: The random variable $y^{n,k+1}$ can be written as:

$$y^{n,k+1} = 1_{\{s^k=n\}}\left[\alpha^k v^k + (1-\alpha^k)y^{n,k}\right] + (1-1_{\{s^k=n\}})y^{n,k}.$$

Using this definition of $y^{n,k+1}$, we get:

$$\begin{aligned}
(y^{n,k+1}-v_n)^2 &= \left\{1_{\{s^k=n\}}\left[\alpha^k(v^k-v_n)+(1-\alpha^k)(y^{n,k}-v_n)\right] + (1-1_{\{s^k=n\}})\left[y^{n,k}-v_n\right]\right\}^2 \\
&= 1_{\{s^k=n\}}\left[(\alpha^k)^2(v^k-v_n)^2 + (1-\alpha^k)^2(y^{n,k}-v_n)^2\right. \\
&\quad \left. + 2\alpha^k(1-\alpha^k)(v^k-v_n)(y^{n,k}-v_n)\right] + (1-1_{\{s^k=n\}})(y^{n,k}-v_n)^2.
\end{aligned}$$

by noting that $(1_H)^2 = 1_H$ and $(1_H)(1-1_H) = 0$ for all $H \in \mathcal{H}$. Taking the expectation of both sides, we get:

$$\begin{aligned}
E\left[(y^{n,k+1}-v_n)^2\right] &= E\left[1_{\{s^k=n\}}(\alpha^k)^2(v^k-v_n)^2\right] + E\left[1_{\{s^k=n\}}(1-\alpha^k)^2(y^{n,k}-v_n)^2\right] \\
&\quad + E\left[1_{\{s^k=n\}}2\alpha^k(1-\alpha^k)(v^k-v_n)(y^{n,k}-v_n)\right] + E\left[(1-1_{\{s^k=n\}})(y^{n,k}-v_n)^2\right].
\end{aligned}$$

We note that:

$$E\left[1_{\{s^k=n\}}2\alpha^k(1-\alpha^k)(v^k-v_n)(y^{n,k}-v_n)\right] = 0$$

since

$$\begin{aligned}
E\left[1_{\{s^k=n\}}2\alpha^k(1-\alpha^k)(v^k-v_n)(y^{n,k}-v_n) \mid s^k=i\right] &= 0 \text{ for } i \neq n \\
E\left[1_{\{s^k=n\}}2\alpha^k(1-\alpha^k)(v^k-v_n)(y^{n,k}-v_n) \mid s^k=n\right] &= 0.
\end{aligned}$$

The second equality follows because given s^k, v^k is independent of v^m and s^m for $m < k$ by A.4, and $y^{n,k}$ is a function of v^m and s^m for $m < k$, and $E\left[v^k | s^k = n\right] = v_n$ by A.3. Then,

$$\begin{aligned}
E\left[(y^{n,k+1}-v_n)^2\right] &= E\left[1_{\{s^k=n\}}(\alpha^k)^2(v^k-v_n)^2\right] \\
&\quad + E\left[1_{\{s^k=n\}}(1+(\alpha^k)^2-2\alpha^k)(y^{n,k}-v_n)^2\right] + E\left[(1-1_{\{s^k=n\}})(y^{n,k}-v_n)^2\right] \\
&= E\left[1_{\{s^k=n\}}(\alpha^k)^2(v^k-v_n)^2\right] \\
&\quad + E\left[1_{\{s^k=n\}}((\alpha^k)^2-2\alpha^k)(y^{n,k}-v_n)^2\right] + E\left[(y^{n,k}-v_n)^2\right].
\end{aligned}$$

Then we get:

$$\begin{aligned}
E\left[(y^{n,k+1}-v_n)^2\right] - E\left[(y^{n,k}-v_n)^2\right] &= E\left[1_{\{s^k=n\}}(\alpha^k)^2(v^k-v_n)^2\right] \\
&\quad + E\left[1_{\{s^k=n\}}((\alpha^k)^2-2\alpha^k)(y^{n,k}-v_n)^2\right].
\end{aligned}$$

Since $\{v^k\}_k$ is uniformly bounded, $\{y^{n,k}\}_k$ is uniformly bounded. Then $\{(y^{n,k}-v_n)^2\}_k$ is also uniformly bounded which implies that $E\left[(y^{n,k}-v_n)^2\right]$ is bounded for all k. If we write the equations above for $k=1,\ldots,N$ and add them side by side, we get:

$$E\left[(y^{n,N+1}-v_n)^2\right]-E\left[(y^{n,1}-v_n)^2\right]=\sum_{k=1}^{N}E\left[1_{\{s^k=n\}}(\alpha^k)^2(v^k-v_n)^2\right]$$
$$+\sum_{k=1}^{N}E\left[1_{\{s^k=n\}}(\alpha^k)^2(y^{n,k}-v_n)^2\right]+\sum_{k=1}^{N}E\left[1_{\{s^k=n\}}(-2\alpha^k)(y^{n,k}-v_n)^2\right].$$

As $N\to\infty$, the left side of the equation is bounded so must be the right side. Since $\sum_k(\alpha^k)^2<\infty$ and v^k are uniformly bounded, the first two terms on the right are bounded. Hence the last term on the right side must be bounded as $N\to\infty$.

$$\infty>\lim_{N\to\infty}\sum_{k=1}^{N}E\left[1_{\{s^k=n\}}(2\alpha^k)(y^{n,k}-v_n)^2\right]=E\sum_{k=1}^{\infty}\left[1_{\{s^k=n\}}(2\alpha^k)(y^{n,k}-v_n)^2\right]$$

by the monotone convergence theorem. Then $\sum_{k=1}^{\infty}\left[1_{\{s^k=n\}}(-2\alpha^k)(y^{n,k}-v_n)^2\right]<\infty$ $a.s.$ But $\sum_k 1_{\{s^k=n\}}\alpha^k=\infty$ $a.s.$ by assumption. Therefore $(y^{n,k}-v_n)^2\to 0$ $a.s.$ which implies that $y^{n,k}\to v_n$ a.s. □

Proof of Lemma 2.3: We start with the first inequality. We use induction. For $k=0$, the result holds by (3). Fix $\omega\in\Omega$ and assume for arbitrary k we have $\hat{v}_j^{n-1,k}(\omega)\leq\hat{v}_j^{n,k}(\omega)$ for all $j\in\{1,\ldots,n-1\}$. We fix $j\in\{1,\ldots,n-1\}$ and consider five cases:

[1] If $1\leq s^k(\omega)<j$: There are four subcases to consider since $1\leq s^k(\omega)<j<n$:

[1.a] $\hat{v}_{s^k}^{n,k+1}(\omega)<\hat{v}_j^{n,k}(\omega)$ and $\hat{v}_{s^k}^{n-1,k+1}(\omega)<\hat{v}_j^{n-1,k}(\omega)$: By the updating procedure we have $\hat{v}_j^{n,k+1}(\omega)=\hat{v}_{s^k}^{n,k+1}(\omega)$ and $\hat{v}_j^{n-1,k+1}(\omega)=\hat{v}_{s^k}^{n-1,k+1}(\omega)$. By the induction hypothesis we have $\hat{v}_{s^k}^{n-1,k}(\omega)\leq\hat{v}_{s^k}^{n,k}(\omega)$. Then $\hat{v}_{s^k}^{n-1,k+1}(\omega)=\alpha^k v^k(\omega)+(1-\alpha^k)\hat{v}_{s^k}^{n-1,k}(\omega)\leq\alpha^k v^k(\omega)+(1-\alpha^k)\hat{v}_{s^k}^{n,k}(\omega)=\hat{v}_{s^k}^{n,k+1}(\omega)$. Putting these three results together, we obtain $\hat{v}_j^{n-1,k+1}(\omega)=\hat{v}_{s^k}^{n-1,k+1}(\omega)\leq\hat{v}_{s^k}^{n,k+1}(\omega)=\hat{v}_j^{n,k+1}(\omega)$.

[1.b] $\hat{v}_{s^k}^{n,k+1}(\omega)<\hat{v}_j^{n,k}(\omega)$ and $\hat{v}_{s^k}^{n-1,k+1}(\omega)\geq\hat{v}_j^{n-1,k}(\omega)$: By the updating procedure we have $\hat{v}_j^{n,k+1}(\omega)=\hat{v}_{s^k}^{n,k+1}(\omega)$ and $\hat{v}_j^{n-1,k+1}(\omega)=\hat{v}_j^{n-1,k}(\omega)$. Similar to [1.a], we have $\hat{v}_{s^k}^{n-1,k+1}(\omega)\leq\hat{v}_{s^k}^{n,k+1}(\omega)$. Putting these three results together we obtain $\hat{v}_j^{n-1,k+1}(\omega)=\hat{v}_j^{n-1,k}(\omega)\leq\hat{v}_{s^k}^{n-1,k+1}(\omega)\leq\hat{v}_{s^k}^{n,k+1}(\omega)=\hat{v}_j^{n,k+1}(\omega)$.

[1.c] $\hat{v}_{s^k}^{n,k+1}(\omega)\geq\hat{v}_j^{n,k}(\omega)$ and $\hat{v}_{s^k}^{n-1,k+1}(\omega)<\hat{v}_j^{n-1,k}(\omega)$: By the updating procedure we have $\hat{v}_j^{n,k+1}(\omega)=\hat{v}_j^{n,k}(\omega)$ and $\hat{v}_j^{n-1,k+1}(\omega)=\hat{v}_{s^k}^{n-1,k+1}(\omega)$. Putting these two results together we obtain $\hat{v}_j^{n-1,k+1}(\omega)=\hat{v}_{s^k}^{n-1,k+1}(\omega)<\hat{v}_j^{n-1,k}(\omega)\leq\hat{v}_j^{n,k}(\omega)=\hat{v}_j^{n,k+1}(\omega)$.

[1.d] $\hat{v}^{n,k+1}_{s^k}(\omega) \geq \hat{v}^{n,k}_j(\omega)$ and $\hat{v}^{n-1,k+1}_{s^k}(\omega) \geq \hat{v}^{n-1,k}_j(\omega)$: By the updating procedure we have $\hat{v}^{n,k+1}_j(\omega) = \hat{v}^{n,k}_j(\omega)$ and $\hat{v}^{n-1,k+1}_j(\omega) = \hat{v}^{n-1,k}_j(\omega)$. Putting these two results together we obtain $\hat{v}^{n-1,k+1}_j(\omega) = \hat{v}^{n-1,k}_j(\omega) \leq \hat{v}^{n,k}_j(\omega) = \hat{v}^{n,k+1}_j(\omega)$.

[2] If $s^k(\omega) = j$: In this case $\hat{v}^{n-1,k+1}_j(\omega) = \hat{v}^{n-1,k+1}_{s^k}(\omega) = \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}^{n-1,k}_{s^k}(\omega) \leq \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}^{n,k}_{s^k}(\omega) = \hat{v}^{n,k+1}_{s^k}(\omega) = \hat{v}^{n,k+1}_j(\omega)$.

[3] If $j < s^k(\omega) \leq n-1$: There are four subcases to consider since $j < s^k(\omega) \leq n-1$:

[3.a] $\hat{v}^{n,k}_j(\omega) < \hat{v}^{n,k+1}_{s^k}(\omega)$ and $\hat{v}^{n-1,k}_j(\omega) < \hat{v}^{n-1,k+1}_{s^k}(\omega)$: By the updating procedure we have $\hat{v}^{n,k+1}_j(\omega) = \hat{v}^{n,k+1}_{s^k}(\omega)$ and $\hat{v}^{n-1,k+1}_j(\omega) = \hat{v}^{n-1,k+1}_{s^k}(\omega)$. Similar to [1.a], by the induction hypothesis we have $\hat{v}^{n-1,k+1}_{s^k}(\omega) = \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}^{n-1,k}_{s^k}(\omega) \leq \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}^{n,k}_{s^k}(\omega) = \hat{v}^{n,k+1}_{s^k}(\omega)$. Putting these three results together we obtain $\hat{v}^{n-1,k+1}_j(\omega) = \hat{v}^{n-1,k+1}_{s^k}(\omega) \leq \hat{v}^{n,k+1}_{s^k}(\omega) = \hat{v}^{n,k+1}_j(\omega)$.

The remaining subcases to consider are:

[3.b] $\hat{v}^{n,k}_j(\omega) < \hat{v}^{n,k+1}_{s^k}(\omega)$ and $\hat{v}^{n-1,k}_j(\omega) \geq \hat{v}^{n-1,k+1}_{s^k}(\omega)$.

[3.c] $\hat{v}^{n,k}_j(\omega) \geq \hat{v}^{n,k+1}_{s^k}(\omega)$ and $\hat{v}^{n-1,k}_j(\omega) < \hat{v}^{n-1,k+1}_{s^k}(\omega)$.

[3.d] $\hat{v}^{n,k}_j(\omega) \geq \hat{v}^{n,k+1}_{s^k}(\omega)$ and $\hat{v}^{n-1,k}_j(\omega) \geq \hat{v}^{n-1,k+1}_{s^k}(\omega)$.

We skip the proofs of these cases since they are similar to cases [1] and [3.a].

[4] If $s^k(\omega) = n$: In this case $\hat{v}^{n-1,k+1}_j(\omega) = \hat{v}^{n-1,k}_j(\omega)$. There are two subcases to consider since $1 \leq j \leq n-1 < n = s^k(\omega)$:

[4.a] $\hat{v}^{n,k}_j(\omega) < \hat{v}^{n,k+1}_{s^k}(\omega)$: By the updating procedure we have $\hat{v}^{n,k+1}_j(\omega) = \hat{v}^{n,k+1}_{s^k}(\omega)$. Then $\hat{v}^{n-1,k+1}_j(\omega) = \hat{v}^{n-1,k}_j(\omega) \leq \hat{v}^{n,k}_j(\omega) < \hat{v}^{n,k+1}_{s^k}(\omega) = \hat{v}^{n,k+1}_j(\omega)$.

[4.b] $\hat{v}^{n,k}_j(\omega) \geq \hat{v}^{n,k+1}_{s^k}(\omega)$: By the updating procedure we have $\hat{v}^{n,k+1}_j(\omega) = \hat{v}^{n,k}_j(\omega)$. Then $\hat{v}^{n-1,k+1}_j(\omega) = \hat{v}^{n-1,k}_j(\omega) \leq \hat{v}^{n,k}_j(\omega) = \hat{v}^{n,k+1}_j(\omega)$.

[5] If $s^k(\omega) \geq n+1$: By the updating procedure we have $\hat{v}^{n,k+1}_j(\omega) = \hat{v}^{n,k}_j(\omega)$, $\hat{v}^{n-1,k+1}_j(\omega) = \hat{v}^{n-1,k}_j(\omega)$. Then $\hat{v}^{n-1,k+1}_j(\omega) = \hat{v}^{n-1,k}_j(\omega) \leq \hat{v}^{n,k}_j(\omega) = \hat{v}^{n,k+1}_j(\omega)$. This shows the first inequality.

We now concentrate on the second inequality. For $k = 0$, the result holds by (7). Fix $\omega \in \Omega$ and assume for arbitrary k we have $y^{n,k}(\omega) \geq \hat{v}^{n,k}_n(\omega)$. We consider three cases:

[1] If $1 \leq s^k(\omega) \leq n-1$: In this case $y^{n,k+1}(\omega) = y^{n,k}(\omega)$ since $s^k(\omega) \neq n$. Then we have two subcases to consider:

[1.a] $\hat{v}_{s^k}^{n,k+1}(\omega) < \hat{v}_n^{n,k}(\omega)$: By the updating procedure we have $\hat{v}_n^{n,k+1}(\omega) = \hat{v}_{s^k}^{n,k+1}(\omega)$. Then $y^{n,k+1}(\omega) = y^{n,k}(\omega) \geq \hat{v}_n^{n,k}(\omega) > \hat{v}_{s^k}^{n,k+1}(\omega) = \hat{v}_n^{n,k+1}(\omega)$.

[1.b] $\hat{v}_{s^k}^{n,k+1}(\omega) \geq \hat{v}_n^{n,k}(\omega)$: By the updating procedure we have $\hat{v}_n^{n,k+1}(\omega) = \hat{v}_n^{n,k}(\omega)$. Then $y^{n,k+1}(\omega) = y^{n,k}(\omega) \geq \hat{v}_n^{n,k}(\omega = \hat{v}_n^{n,k+1}(\omega)$.

[2] If $s^k(\omega) = n$: In this case $y^{n,k+1}(\omega) = \alpha^k v^k(\omega) + (1-\alpha^k) y^{n,k}(\omega) \geq \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_n^{n,k}(\omega) = \hat{v}_n^{n,k}(\omega)$.

[3] If $s^k(\omega) \geq n+1$: In this case $y^{n,k+1}(\omega) = y^{n,k}(\omega) \geq \hat{v}_n^{n,k}(\omega) = \hat{v}_n^{n,k+1}(\omega)$. □

Proof of Lemma 2.6: Using (4), we can write:

$$U_{j,n+1}^{n+1,k} = \left\{\omega \in \Omega : s^k(\omega) = j, \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_j^{n+1,k}(\omega) < \hat{v}_{n+1}^{n+1,k}(\omega)\right\}$$

for $j \in \{1, \ldots, n\}$. In the following, the second and third lines are implied by lemma 2.3 and the fourth line is implied by the updating process of $\hat{v}_j^{n,k}$:

$$\begin{aligned} U_{j,n+1}^{n+1,k} &= \left\{\omega \in \Omega : s^k(\omega) = j, \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_j^{n+1,k}(\omega) < \hat{v}_{n+1}^{n+1,k}(\omega)\right\} \\ &\subset \left\{\omega \in \Omega : s^k(\omega) = j, \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_j^{n,k}(\omega) < \hat{v}_{n+1}^{n+1,k}(\omega)\right\} \\ &\subset \left\{\omega \in \Omega : s^k(\omega) = j, \alpha^k v^k(\omega) + (1-\alpha^k)\hat{v}_j^{n,k}(\omega) < \hat{v}^{n+1,k}(\omega)\right\} \\ &= \left\{\omega \in \Omega : s^k(\omega) = j, \hat{v}_j^{n,k+1}(\omega) < y^{n+1,k}(\omega)\right\} \\ &\subset \left\{\omega \in \Omega : \hat{v}_j^{n,k+1}(\omega) < y^{n+1,k}(\omega)\right\}. \end{aligned}$$

Thus $\sum_k 1_{\{\omega \in U_{j,n+1}^{n+1,k}\}} \leq \sum_k 1_{\{\hat{v}_j^{n,k+1}(\omega) < y^{n+1,k}(\omega)\}}$ for all ω. By the assumption that $\hat{v}_j^{n,k} \to v_j$ *a.s.* for all $j \in \{1, \ldots, n\}$ and lemma 2.1, we have $lim_{k\to\infty}\hat{v}_j^{n,k+1}(\omega) - y^{n+1,k}(\omega) = v_j - v_{n+1}$ for *a.e.* ω. Then for any $\epsilon > 0$, $\sum_k 1_{(\epsilon,\infty)} \circ |(\hat{v}_j^{n,k+1}(\omega) - y^{n+1,k}(\omega)) - (v_j - v_{n+1})| < \infty$ for *a.e.* ω. Pick $\epsilon = v_j - v_{n+1} > 0$:

$$\begin{aligned} \sum_k 1_{\{\hat{v}_j^{n,k+1} < y^{n+1,k}\}} &= \sum_k 1_{\{(y^{n+1,k} - \hat{v}_j^{n,k+1}) - (v_{n+1} - v_j) > v_j - v_{n+1}\}} \\ &\leq \sum_k 1_{\{|(y^{n+1,k} - \hat{v}_j^{n,k+1}) - (v_{n+1} - v_j)| > v_j - v_{n+1}\}} \\ &= \sum_k 1_{(v_j - v_{n+1},\infty)} \circ |(\hat{v}_j^{n,k+1} - y^{n+1,k}) - (v_j - v_{n+1})| < \infty \quad a.s. \end{aligned}$$

So $\sum_k 1_{U_{j,n+1}^{n+1,k}} \leq \sum_k 1_{\{\hat{v}_j^{n,k+1} < y^{n+1,k}\}} < \infty$ *a.s.*

The proof for $\sum_k 1_{U_{n+1,j}^{n+1,k}} < \infty$ *a.s.* follows the same lines. □

Proof of Proposition 2.1: In order to complete the proof, we need to show that $\hat{v}_{n+1}^{n+1,k} \to v_{n+1}$ *a.s.* This part of the proof is similar to the proof of lemma 2.5. $y^{n+1,k+1}(\omega)$ can be written as:

$$y^{n+1,k+1}(\omega) \;=\; 1_{\{s^k(\omega)=n+1\}}\left[\alpha^k v^k(\omega) + (1-\alpha^k) y^{n+1,k}(\omega)\right] + (1 - 1_{\{s^k(\omega)=n+1\}})\; y^{n+1,k}(\omega).$$

On the other hand using (5), $\hat{v}_{n+1}^{n+1,k+1}(\omega)$ can be written as:

$$\begin{aligned}\hat{v}_{n+1}^{n+1,k+1}(\omega) &= 1_{\{s^k(\omega)=n+1\}}\left[\alpha^k v^k(\omega)+(1-\alpha^k)\hat{v}_{n+1}^{n+1,k}(\omega)\right]+\sum_{i=1}^{n+1}1_{\{\omega\in U_{i,n+1}^{n+1,k}\}}\hat{v}_i^{n+1,k+1}(\omega)\\ &\quad +(1-1_{\{s^k(\omega)=n+1\}}-\sum_{i=1}^{n+1}1_{\{\omega\in U_{i,n+1}^{n+1,k}\}})\hat{v}_{n+1}^{n+1,k}(\omega).\end{aligned}$$

If we subtract the two equations above side by side, we get:

$$\begin{aligned}y^{n+1,k+1}(\omega)-\hat{v}_{n+1}^{n+1,k+1}(\omega) &= \left[1_{\{s^k(\omega)=n+1\}}(1-\alpha^k)+(1-1_{\{s^k(\omega)=n+1\}})\right]\left[y^{n+1,k}(\omega)-\hat{v}_{n+1}^{n+1,k}(\omega)\right]\\ &\quad -\sum_{i=1}^{n+1}1_{\{\omega\in U_{i,n+1}^{n+1,k}\}}\left[\hat{v}_i^{n+1,k+1}(\omega)-\hat{v}_{n+1}^{n+1,k}(\omega)\right].\end{aligned}$$

By lemma 2.6, for all $i\in\{1,\ldots,n\}$ and $a.e.\ \omega$, there exists a finite $N(\omega)$ such that $1_{\{\omega\in U_{i,n+1}^{n+1,k}\}}=0$ for all $k\geq N(\omega)$. Then for $k\geq N(\omega)$:

$$\begin{aligned}y^{n+1,k+1}(\omega)-\hat{v}_{n+1}^{n+1,k+1}(\omega) &= \left[1_{\{s^k(\omega)=n+1\}}(1-\alpha^k)+(1-1_{\{s^k(\omega)=n+1\}})\right]\left[y^{n+1,k}(\omega)-\hat{v}_{n+1}^{n+1,k}(\omega)\right]\\ \left|y^{n+1,k+1}(\omega)-\hat{v}_{n+1}^{n+1,k+1}(\omega)\right| &= \left[1_{\{s^k(\omega)=n+1\}}(1-\alpha^k)+(1-1_{\{s^k(\omega)=n+1\}})\right]\left|y^{n+1,k}(\omega)-\hat{v}_{n+1}^{n+1,k}(\omega)\right|.\end{aligned}$$

This implies that

$$\left|y^{n+1,k+1}(\omega)-\hat{v}_{n+1}^{n+1,k+1}(\omega)\right|-\left|y^{n+1,k}(\omega)-\hat{v}_{n+1}^{n+1,k}(\omega)\right|=-1_{\{s^k(\omega)=n+1\}}\alpha^k\left|y^{n+1,k}(\omega)-\hat{v}_{n+1}^{n+1,k}(\omega)\right|.$$

If we write the equation above for $k=N(\omega)$ to $k=K>N(\omega)$ and add them side by side, we get:

$$\begin{aligned}&\left|y^{n+1,K+1}(\omega)-\hat{v}_{n+1}^{n+1,K+1}(\omega)\right|-\left|y^{n+1,N(\omega)}(\omega)-\hat{v}_{n+1}^{n+1,N(\omega)}(\omega)\right|=\\ &\qquad -\sum_{k=N(\omega)}^{K}1_{\{s^k(\omega)=n+1\}}\alpha^k\left|y^{n+1,k}(\omega)-\hat{v}_{n+1}^{n+1,k}(\omega)\right|.\end{aligned}$$

Since v^k are uniformly bounded, if we take the limit of the left hand side as $K\to\infty$, the left side is bounded. So must be the right side. But by assumption $\sum_{k=N(\omega)}^{\infty}1_{\{s^k(\omega)=n+1\}}\alpha^k=\infty$ for $a.e.\ \omega$ since $N(\omega)$ is finite for $a.e.\ \omega$. Then $\lim_{k\to\infty}\left|y^{n+1,k}(\omega)-\hat{v}_{n+1}^{n+1,k}(\omega)\right|\to 0$ for $a.e.\ \omega$. $y^{n+1,k}\to v_{n+1}$ $a.s.$ by lemma 2.1, hence $\hat{v}_{n+1}^{n+1,k}\to v_{n+1}$ $a.s.$ $\square$