Corrections for

DYNAMIC PROGRAMMING AND OPTIMAL CONTROL: 4TH and EARLIER EDITIONS

by Dimitri P. Bertsekas

Athena Scientific

Last Updated: 4/19/25

VOLUME 1 - 4TH EDITION, 2ND PRINTING, 2020

p. 5 Correct Fig. 1.1.2 so that the system equation reads

$$x_{k+1} = x_k + u_k - w_k$$

and the cost per stage reads

$$r(x_k) + cu_k$$

p. 47 Change the last equation to

$$\begin{aligned}
J_k^\epsilon(x_k) &= \mathop{E}_{w_k}\big\{g_k\big(x_k,\mu_k^\epsilon(x_k),w_k\big) + J_{k+1}^\epsilon\big(f_k\big(x_k,\mu_k^\epsilon(x_k),w_k\big)\big)\big\} \\
&\le \mathop{E}_{w_k}\big\{g_k\big(x_k,\mu_k^\epsilon(x_k),w_k\big) + J_{k+1}\big(f_k\big(x_k,\mu_k^\epsilon(x_k),w_k\big)\big)\big\} + (N-k-1)\epsilon \\
&\le J_k(x_k) + \epsilon + (N-k-1)\epsilon \\
&= \min_{u_k\in U(x_k)} \mathop{E}_{w_k}\big\{g_k\big(x_k,u_k,w_k\big) + J_{k+1}\big(f_k\big(x_k,u_k,w_k\big)\big)\big\} + (N-k)\epsilon \\
&\le \mathop{E}_{w_k}\big\{g_k\big(x_k,\mu_k(x_k),w_k\big) + J_{\pi^{k+1}}\big(f_k\big(x_k,\mu_k(x_k),w_k\big)\big)\big\} + (N-k)\epsilon \\
&= J_{\pi^k}(x_k) + (N-k)\epsilon,
\end{aligned}$$

p. 66 (-4) Change the equation to

$$V_k(y_k) = E_{w_k}\bigg\{J_{k+1}\big(h_k(y_k,w_k),w_k\big)\bigg\}.$$

p. 77 (-2) Change "HHM" to "HMM"

p. 94 (+5) Change $\underline{f}_Y$ to $\underline{f}_Y - \underline{f}_X$

p. 221 (-13) Change "Chapter 3" to "Chapter 4"

p. 294 (-4) Change "paper" to "book"

p. 430 (+4) Replace the part

Assuming that $\tilde{J}^*$ has the required differentiability properties, we expand it into a first order Taylor series around $(k\delta, x)$, obtaining

$$\tilde{J}^*\big((k+1)\cdot\delta, x + f(x,u)\cdot\delta\big) = \tilde{J}^*(k\delta, x) + \nabla_t \tilde{J}^*(k\delta, x)\cdot\delta + \nabla_x \tilde{J}^*(k\delta, x)' f(x,u)\cdot\delta + o(\delta),$$

where $o(\delta)$ represents second order terms satisfying $\lim_{\delta\to 0} o(\delta)/\delta = 0$, ∇_t denotes partial derivative with respect to t, and ∇_x denotes the n-dimensional (column) vector of partial derivatives with respect to x. Substituting in the DP equation, we obtain

$$\tilde{J}^*(k\delta, x) = \min_{u\in U}\big[g(x,u)\cdot\delta + \tilde{J}^*(k\delta, x) + \nabla_t \tilde{J}^*(k\delta, x)\cdot\delta + \nabla_x \tilde{J}^*(k\delta, x)' f(x,u)\cdot\delta + o(\delta)\big].$$

Canceling $\tilde{J}^*(k\delta, x)$ from both sides, dividing by δ, and taking the limit as $\delta \to 0$, while assuming that the discrete-time cost-to-go function yields in the limit its continuous-time counterpart,

$$\lim_{k\to\infty,\,\delta\to 0,\,k\delta=t} \tilde{J}^*(k\delta, x) = J^*(t,x), \qquad \text{for all } t, x,$$

we obtain the following equation for the cost-to-go function $J^*(t,x)$:

$$0 = \min_{u\in U}\big[g(x,u) + \nabla_t J^*(t,x) + \nabla_x J^*(t,x)' f(x,u)\big], \qquad \text{for all } t, x,$$

with the boundary condition $J^*(T,x) = h(x)$.

with the following:

Assuming that $\tilde{J}^*$ has the required differentiability properties, we expand it into a first order Taylor series around $\big((k+1)\delta, x\big)$, obtaining

$$\tilde{J}^*\big((k+1)\cdot\delta, x + f(x,u)\cdot\delta\big) = \tilde{J}^*\big((k+1)\delta, x\big) + \nabla_x \tilde{J}^*\big((k+1)\delta, x\big)' f(x,u)\cdot\delta + o(\delta),$$

where $o(\delta)$ represents second order terms satisfying $\lim_{\delta\to 0} o(\delta)/\delta = 0$, and ∇_x denotes the n-dimensional (column) vector of partial derivatives with respect to x. Substituting in the DP equation, we obtain

$$\tilde{J}^*(k\delta, x) = \min_{u\in U}\Big[g(x,u)\cdot\delta + \tilde{J}^*\big((k+1)\delta, x\big) + \nabla_x \tilde{J}^*\big((k+1)\delta, x\big)' f(x,u)\cdot\delta + o(\delta)\Big].$$

By dividing with δ, and by introducing the "discretized" partial derivative with respect to t, given by

$$\widehat{\nabla}_t \tilde{J}^*(k\delta, x) = \frac{\tilde{J}^*\big((k+1)\delta, x\big) - \tilde{J}^*\big(k\delta, x\big)}{\delta},$$

we obtain

$$0 = \min_{u \in U} \left[g(x,u) + \widehat{\nabla}_t \tilde{J}^*(k\delta, x) + \nabla_x \tilde{J}^*\big((k+1)\delta, x\big)' f(x,u) + \frac{o(\delta)}{\delta} \right].$$

Taking the limit as $\delta \to 0$, while assuming that

$$\lim_{k\to\infty,\, \delta\to 0,\, k\delta = t} \widehat{\nabla}_t \tilde{J}^*(k\delta, x) = \nabla_t J^*(t,x), \qquad \text{for all } t, x,$$

$$\lim_{k\to\infty,\, \delta\to 0,\, k\delta = t} \nabla_x \tilde{J}^*\big((k+1)\delta, x\big) = \nabla_x J^*(t,x), \qquad \text{for all } t, x,$$

we obtain the following equation for the cost-to-go function $J^*(t,x)$:

$$0 = \min_{u \in U} \big[g(x,u) + \nabla_t J^*(t,x) + \nabla_x J^*(t,x)' f(x,u) \big], \qquad \text{for all } t, x,$$

with the boundary condition $J^*(T,x) = h(x)$.

p. 490 (+4) Change $x(\cdot) : \Re^m \to \Re^m$ to $x(\cdot) : \Re^m \to \Re^n$

p. 490 (+12) Change $x(\cdot) : \Re^n \to \Re^m$ to $x(\cdot) : \Re^m \to \Re^n$

VOLUME 1 - 4TH EDITION, 1ST PRINTING, 2017

p. 5 Correct Fig. 1.1.2 so that the system equation reads

$$x_{k+1} = x_k + u_k - w_k$$

and the cost per stage reads

$$r(x_k) + cu_k$$

p. 47 Change the last equation to

$$\begin{aligned}
J_k^\epsilon(x_k) &= \mathop{E}_{w_k}\big\{g_k\big(x_k, \mu_k^\epsilon(x_k), w_k\big) + J_{k+1}^\epsilon\big(f_k\big(x_k, \mu_k^\epsilon(x_k), w_k\big)\big)\big\} \\
&\le \mathop{E}_{w_k}\big\{g_k\big(x_k, \mu_k^\epsilon(x_k), w_k\big) + J_{k+1}\big(f_k\big(x_k, \mu_k^\epsilon(x_k), w_k\big)\big)\big\} + (N-k-1)\epsilon \\
&\le J_k(x_k) + \epsilon + (N-k-1)\epsilon \\
&= \min_{u_k \in U(x_k)} \mathop{E}_{w_k}\big\{g_k\big(x_k, u_k, w_k\big) + J_{k+1}\big(f_k\big(x_k, u_k, w_k\big)\big)\big\} + (N-k)\epsilon \\
&\le \mathop{E}_{w_k}\big\{g_k\big(x_k, \mu_k(x_k), w_k\big) + J_{\pi^{k+1}}\big(f_k\big(x_k, \mu_k(x_k), w_k\big)\big)\big\} + (N-k)\epsilon \\
&= J_{\pi^k}(x_k) + (N-k)\epsilon,
\end{aligned}$$

p. 66 (-4) Change the equation to

$$V_k(y_k) = E_{w_k}\bigg\{J_{k+1}\big(h_k(y_k, w_k), w_k\big)\bigg\}.$$

p. 295 (-9) Change $\tilde{J}_k$ to $\tilde{J}_{k+1}$

p. 301 (-9) Change x_k to x_{N-1}

p. 324 (-4) Change "Example 3.3.1" to "Example 6.3.1"

p. 339 (+15) Change $i = 1, \ldots, q$ to $s = 1, \ldots, q$

p. 342 (+8) Change "approximations" to "approximations and simulation-based implementations"

p. 430 (+4) Replace the part

Assuming that $\tilde{J}^*$ has the required differentiability properties, we expand it into a first order Taylor series around $(k\delta, x)$, obtaining

$$\begin{aligned}
\tilde{J}^*\big((k+1)\cdot\delta, x + f(x,u)\cdot\delta\big) = \tilde{J}^*(k\delta, x) &+ \nabla_t \tilde{J}^*(k\delta, x)\cdot\delta \\
&+ \nabla_x \tilde{J}^*(k\delta, x)' f(x,u)\cdot\delta + o(\delta),
\end{aligned}$$

where $o(\delta)$ represents second order terms satisfying $\lim_{\delta\to 0} o(\delta)/\delta = 0$, ∇_t denotes partial derivative with respect to t, and ∇_x denotes the n-dimensional (column) vector of partial derivatives with respect to x. Substituting in the DP equation, we obtain

$$\begin{aligned}\tilde J^*(k\delta,x) = \min_{u\in U}\big[g(x,u)\cdot\delta + \tilde J^*(k\delta,x) + \nabla_t \tilde J^*(k\delta,x)\cdot\delta \\ + \nabla_x \tilde J^*(k\delta,x)' f(x,u)\cdot\delta + o(\delta)\big].\end{aligned}$$

Canceling $\tilde J^*(k\delta,x)$ from both sides, dividing by δ, and taking the limit as $\delta\to 0$, while assuming that the discrete-time cost-to-go function yields in the limit its continuous-time counterpart,

$$\lim_{k\to\infty,\,\delta\to 0,\,k\delta=t} \tilde J^*(k\delta,x) = J^*(t,x), \qquad \text{for all } t,x,$$

we obtain the following equation for the cost-to-go function $J^*(t,x)$:

$$0 = \min_{u\in U}\big[g(x,u) + \nabla_t J^*(t,x) + \nabla_x J^*(t,x)' f(x,u)\big], \qquad \text{for all } t,x,$$

with the boundary condition $J^*(T,x) = h(x)$.

with the following:

Assuming that $\tilde J^*$ has the required differentiability properties, we expand it into a first order Taylor series around $\big((k+1)\delta,x\big)$, obtaining

$$\begin{aligned}\tilde J^*\big((k+1)\cdot\delta, x + f(x,u)\cdot\delta\big) = \tilde J^*\big((k+1)\delta,x\big) \\ + \nabla_x \tilde J^*\big((k+1)\delta,x\big)' f(x,u)\cdot\delta + o(\delta),\end{aligned}$$

where $o(\delta)$ represents second order terms satisfying $\lim_{\delta\to 0} o(\delta)/\delta = 0$, and ∇_x denotes the n-dimensional (column) vector of partial derivatives with respect to x. Substituting in the DP equation, we obtain

$$\begin{aligned}\tilde J^*(k\delta,x) = \min_{u\in U}\Big[g(x,u)\cdot\delta + \tilde J^*\big((k+1)\delta,x\big) \\ + \nabla_x \tilde J^*\big((k+1)\delta,x\big)' f(x,u)\cdot\delta + o(\delta)\Big].\end{aligned}$$

By dividing with δ, and by introducing the "discretized" partial derivative with respect to t, given by

$$\hat\nabla_t \tilde J^*(k\delta,x) = \frac{\tilde J^*\big((k+1)\delta,x\big) - \tilde J^*\big(k\delta,x\big)}{\delta},$$

we obtain

$$0 = \min_{u\in U}\left[g(x,u) + \hat\nabla_t \tilde J^*(k\delta,x) + \nabla_x \tilde J^*\big((k+1)\delta,x\big)' f(x,u) + \frac{o(\delta)}{\delta}\right].$$

Taking the limit as $\delta \to 0$, while assuming that

$$\lim_{k\to\infty,\, \delta\to 0,\, k\delta=t} \hat{\nabla}_t \tilde{J}^*(k\delta, x) = \nabla_t J^*(t,x), \qquad \text{for all } t, x,$$

$$\lim_{k\to\infty,\, \delta\to 0,\, k\delta=t} \nabla_x \tilde{J}^*\big((k+1)\delta, x\big) = \nabla_x J^*(t,x), \qquad \text{for all } t, x,$$

we obtain the following equation for the cost-to-go function $J^*(t,x)$:

$$0 = \min_{u\in U}\big[g(x,u) + \nabla_t J^*(t,x) + \nabla_x J^*(t,x)' f(x,u)\big], \qquad \text{for all } t, x,$$

with the boundary condition $J^*(T,x) = h(x)$.

p. 490 (+4) Change $x(\cdot) : \Re^m \to \Re^m$ to $x(\cdot) : \Re^m \to \Re^n$

p. 490 (+12) Change $x(\cdot) : \Re^n \to \Re^m$ to $x(\cdot) : \Re^m \to \Re^n$

VOLUME 2 - 4TH EDITION, 2ND PRINTING, 2018

p. 14 (+12) Change (x,u,w) to (x,u).

p. 14 (+13) Change $\big|E\{g(x,u)\}\big| \le b$ to $\big|E\{g(x,u,w)\}\big| \le b$.

p. 258 (+14) Change line

$$\big|g_k(x_k,u_k,w_k)\big| \le M, \qquad \text{for all } (x_k,u_k,w_k) \in X_k \times U_k \times W_k,$$

to

$$\Big|E\big\{g_k(x_k,u_k,w_k)\big\}\Big| \le M, \qquad \text{for all } (x_k,u_k) \in X_k \times U_k,$$

p. 258 (+17) Change line

$$0 \le g_k(x_k,u_k,w_k), \qquad \text{for all } (x_k,u_k,w_k) \in X_k \times U_k \times W_k,$$

to

$$0 \le E\big\{g_k(x_k,u_k,w_k)\big\}, \qquad \text{for all } (x_k,u_k) \in X_k \times U_k,$$

p. 258 (+19) Change line

$$g_k(x_k,u_k,w_k) \le 0, \qquad \text{for all } (x_k,u_k,w_k) \in X_k \times U_k \times W_k,$$

to

$$E\big\{g_k(x_k,u_k,w_k)\big\} \le 0, \qquad \text{for all } (x_k,u_k) \in X_k \times U_k,$$

p. 435 (-11) Change Exercise 6.7 to Exercise 6.8.

p. 457 (-11) Change "weighted sup-norm" to "weighted Euclidean norm".

p. 472 Change the third equation from the bottom to

$$C_\mu = \Phi'\Xi_\mu(1-\alpha P_\mu)\Phi = (\,1 \quad 2\,)\begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}\begin{pmatrix} 1 & -\alpha \\ -a & 1 \end{pmatrix}\begin{pmatrix} 1 \\ 2 \end{pmatrix} = \frac{5-4\alpha}{2},$$

p. 472 Change the second equation from the bottom to

$$d_\mu = \Phi'\Xi_\mu g_\mu = (\,1 \quad 2\,)\begin{pmatrix} 1/2 & 0 \\ 0 & 1/2 \end{pmatrix}\begin{pmatrix} 0 \\ 0 \end{pmatrix} = 0,$$

p. 473 Change the second equation from the top to

$$d_{\mu^*} = \Phi'\Xi_{\mu^*}g_{\mu^*} = (\,1 \quad 2\,)\begin{pmatrix} \frac{1}{2-p} & 0 \\ 0 & \frac{1-p}{2-p} \end{pmatrix}\begin{pmatrix} pc \\ 0 \end{pmatrix} = \frac{pc}{2-p},$$

p. 473 Change the third equation from the top to

$$r_{\mu^*} = \frac{pc}{5-4p-\alpha(4-3p)}.$$

p. 492 (+6) Change $C_k r$ to $C_k r_k$.

VOLUME 2 - 4TH EDITION, 1ST PRINTING, 2012

p. 63 (+9) Change equation to

$$\lim_{k\to\infty} \|J_\mu - T_\mu^k J\| = 0, \qquad \lim_{k\to\infty} \|J^* - T^k J\| = 0.$$

p. 68 (-9) Change "Applying Eq. (1.60) with J and TJ and replaced by J' and J, respectively" to "Applying Eq. (1.60) with J' and J replaced by J and TJ, respectively"

p. 87 (-7) Change "$\alpha c_1 \leq c_2$" to "$\alpha \underline{c}_1 \leq \underline{c}_2$"

p. 125 (-9 and -11) Change "Section 1.6.1" to "Section 1.6.2"

p. 135 (-5, -4, and -3) The calculation in these three lines has a flaw, and in the final result $\overline{\mu}$ should be replaced by μ. Change these three lines to the following:

$$\|T\bar{J} - T_\mu \bar{J}\| \leq \|T\bar{J} - T_{\overline{\mu}}\bar{J}\| + \|T_{\overline{\mu}}\bar{J} - T_\mu \bar{J}\| \leq \epsilon + \zeta,$$

and by replacing ϵ with $\epsilon + \zeta$ and $\overline{\mu}$ with μ in Eq. (2.79), we obtain

$$\|J_\mu - J^*\| \leq \frac{\epsilon + \zeta + 2\alpha\delta}{1-\alpha}.$$

p. 136 (+17) Change "$\mu \leq \mathcal{M}$" to "$\mu \in \mathcal{M}$"

p. 141 (-7) Change "$J \in S(0)$" to "$J^0 \in S(0)$"

p. 142 (20) Change "$J \in S(0)$" to "$J^0 \in S(0)$"

p. 147 (-7) Change "$\forall\ \mu \in \mathcal{M}$" to "for some $\mu \in \mathcal{M}$"

p. 163 (1) Part (a) of Exercise 2.9 requires additional assumptions, such as $p_{ij} = 0$ for all i and $j \neq i, i+1$. See Exercise 7.8 of Vol. I.

p. 230 (-5) Change "X, U, and W" to "X and U"

p. 230 (-1) Change "$g(i,u) + \sum_{j=1}^n p_{ij}(u) z_j$" to "$\sum_{j=1}^n p_{ij}(u)\big(g(i,u,j) + z_j\big)$"

p. 231 (+4) Change "$g(i,u) + \sum_{j=1}^n p_{ij}(u) z_j$" to "$\sum_{j=1}^n p_{ij}(u)\big(g(i,u,j) + z_j\big)$"

p. 237 (-7) Change "$t\big(f_c(x,w)\big)$" to "$s\big(f_c(x,w)\big)$"

p. 237 (Last paragraph) For a one-step lookahead policy to be optimal it is not enough that condition (4.24) holds. It is also necessary that the monotonically nonincreasing) sequence $(T^k s)(x)$ converges to $J^*(x)$ for all

$x \in X$. This is needed for the argument of Section 4.4, Vol. I, to go through, but may not be true under Assumption P. It is, however, true for the case of Example 4.4.1 of the next page, so the analysis of that example is correct.

p. 412 (+18) Change "$\bar{b} = E\{A(w)\}$" to "$\bar{b} = E\{b(w)\}$"

p. 415 (+14) Change "probabilities and stage costs" to "probabilities"

p. 435 (-17) Change "implementation" to "implementations"

p. 436 (-10) Change "to to" to "to"

p. 461 (+13) Change "product" to "sum of the products"

p. 461 (-12) Change "$w_\ell(i)(T^{(w)}J)(i)$" to "$w_\ell(i)(T^\ell J)(i)$"

p. 491 (-10) Change "$\tilde{J}_k(i) \to J^*(i)$" to "$\tilde{J}_0(i) \to J^*(i)$"

p. 492 (-11) Change "i belongs" to "j belongs"

p. 508 (+15) Change "issue" to "issues"

p. 559 (+6) Change "$C^{-1}d_k$" to "$C^{-1}d$"

p. 567 (-20) Change "row" to "column"

p. 575 (+11) Change "γP" to "$\gamma \Pi P$"

p. 641 (-5) Change "assume reader" to "assume a reader"

p. 654 (+13) Change "$\overline{\mu}_k(u_k)$" to "$\overline{\mu}_k(x_k)$" (twice)

p. 663 (+13) Insert additional reference

[Bla65] Blackwell, D., 1965. "Positive Dynamic Programming," Proc. Fifth Berkeley Symposium Math. Statistics and Probability, pp. 415-418.

VOLUME 1 - 3RD EDITION

p. vi (-16) Change title of Chapter 6 to "Approximate Dynamic Programming"

p. 7 (+19) Change "... after operation B ..." to "... after operation C ..."

p. 52 (+5) Change "$J_0(1) = 2.67$ and $J_0(2) = 2.608$" to "$J_0(1) = 2.7$ and $J_0(2) = 2.818$"

p. 99 (+1) Change "**Figure 2.4.1**" to "**Figure 2.4.2**"

p. 106 (-15) Add footnote: "Our definition of piecewise continuous control functions assumes a finite number of points of discontinuity within $[0, T]$. The solution $x(t)$ of the differential system (3.1) is then continuous for $t \in [0, T]$, and differentiable at all points $t \in [0, T]$ where $u(t)$ is continuous. Thus, formally, for given piecewise continuous $u(t)$, a solution of Eq. (3.1) is a continuous function $x(t)$ which is differentiable at all points of continuity of $u(t)$ and satisfies Eq. (3.1) at these points."

p. 153 (-12 Relabel "Proposition 4.4.1" to "Proposition 4.1.1"

p. 171 (-3) Change "$e^{-x/a}$" to "$-e^{-x/a}$"

p. 172 (-8) Change "Using relation (4.31)," to "Using relations (4.30) and (4.31),"

p. 185 (+1) Technically Example 4.4.1 is not a time-homogeneous problem. However, because the DP equation is time-homogeneous the analysis of pp. 183-184 on one-step lookahead rule applies (the proof only uses time-homogeneity of the DP mapping).

p. 213 (+9) This problem is flawed as stated. Replace the statement with the following:

4.27

Consider the quiz contest problem of Example 5.1, where the questions are partitioned in M groups, and there is an order constraint that all the questions in group m must be answered before the questions in group $m+1$ can be answered. Show that an optimal list can be constructed by ordering the questions within each group in decreasing order of $p_i R_i/(1-p_i)$. Consider also the problem of optimally ordering the groups in addition to optimally ordering the questions within each group. Show that it is optimal to answer groups in order of decreasing $W/(1-P)$, where for a given group, W is the expected reward obtained by answering only the questions of that group and in optimal order, and P is the probability of answering all the questions of the group correctly.

p. 214 (+2) Change "$[\beta, \beta]$" to "$[-\beta, \beta]$"

p. 222 (+7) Change "and for $k = 0, 1, \ldots, N-2$,

$$J_k(I_k) = \min_{u_k \in U_k} \left[\mathop{E}_{x_k,\, w_k,\, z_{k+1}} \left\{ g_k(x_k, u_k, w_k) + J_{k+1}(I_k, z_{k+1}, u_k) \mid I_k, u_k \right\} \right]. \tag{5.5}$$

to "and for $k = 0, 1, \ldots, N-2$,

$$J_k(I_k) = \min_{u_k \in U_k} \left[\mathop{E}_{z_{k+1}} \left\{ \tilde{g}_k(I_k, u_k) + J_{k+1}(I_k, z_{k+1}, u_k) \mid I_k, u_k \right\} \right]$$

or equivalently

$$J_k(I_k) = \min_{u_k \in U_k} \left[\mathop{E}_{x_k,\, w_k,\, z_{k+1}} \left\{ g_k(x_k, u_k, w_k) + J_{k+1}(I_k, z_{k+1}, u_k) \mid I_k, u_k \right\} \right]. \tag{5.5}$$

p. 265 (+7 and +8) Change "$\overline{\alpha}_k$" to "$\overline{\alpha}$"

p. 268 (+6) Change "$(1-p)L_0, pL_0$" to "$(1-p)L_0, pL_1$"

p. 270 (-5) Change "... Exercise 5.6 ..." to "... Exercise 5.3 ..."

p. 327 (+21) Change "relies provides" to "provides"

p. 331 (-8) Change "$-\infty$" to "∞"

p. 331 (-9) Change "∞" to "$-\infty$"

p. 387 (+20) Change "... all which ..." to "... all of which ..."

p. 393 (+8) Change "$\lim_{|x|\to 1} g(x) = \infty$" to "$\lim_{|x|\to\infty} g(x) = \infty$"

p. 406 (+16) Change "... no more that" to "... no more than"

p. 426 (+6) Change "... λ^* is the same" to "is a constant λ^*"

p. 428 (-16) Change "... for all i and k." to "... for all i."

p. 431 (-9) Change "... tentative backed-up score $TBS(n)$ of position n to ∞ if it is White's turn to move at n and to $-\infty$ if it is Black's turn ..." to "...tentative backed-up score $TBS(n)$ of position n to $-\infty$ if it is White's turn to move at n and to ∞ if it is Black's turn ..."

p. 443 (+10, +21, +24) Change "N_{nn}" to "T_{nn}"

p. 446 (1) In this problem we do not assume the 10-yard rule, which allows for a return to a first down, once progress by at least 10 yards has been made within four downs. Thus the drive will last for at most four downs. Alternatively, you may assume the 10-yard rule. The problem will be similar, but the number of states will be larger.

p. 453 (1) Change "decreases monotonically with i" to "is monotonically nondecreasing with i"

p. 453 (-16) Change "$1 - p_A p_B p_C$" to "$p_A p_B p_C$"

p. 453 (-15) Change "decreasing" to "nondecreasing"

p. 453 (-8) Change "$1 + p_A p_B$" to "$1 + p_A + p_A p_B$"

p. 467 (+4) Change "$\partial x^i \partial x^j$" to "$\partial x_i \partial x_j$"

p. 479 (-14) Change "... Prop. A1 of Appendix A in Vol. II." to "... Section 4.1 of Vol. II."

p. 512 (-3) Change "... with conceptually convenient ..." to "... with the conceptually convenient ..."

VOLUME 2 - 3RD EDITION

p. 51 (+12) Delete the phrase "where $\tilde{S}$ ... $S = (1, \ldots, n)$."

p. 57 (+2) Change "satisfying $\|y\| \geq 0$ for all $y \in Y$, $\|y\| = 0$ if and only if $y = 0$" to "satisfying for all $y \in Y$, $\|y\| \geq 0$, $\|y\| = 0$ if and only if $y = 0$, $\|ay\| = |a|\|y\|$ for all scalars a,"

p. 63 (+19) Change "$\to 0$" to "$= 0$"

p. 63 (+21) Change "$\to 0$" to "$= 0$"

p. 119 (-8) Change "a proper policy" to "an optimal proper policy"

p. 120 (-2) Change "of the multistage policy iteration algorithm discussed in Section 2.3.3." to "of a multistage policy iteration algorithm."

p. 198 (-10) Change Prop. 4.1.9 to Prop. 4.2.1

p. 274 (+21) Change Section 6.4.2 to Section 4.6.2

p. 341 (+14) Change $\|J - \hat{J}\|$ to $\|J - \hat{J}\|_v$

p. 341 (-4) Change $\sum_{i=1}^n$ to $\sum_{j=1}^n$

p. 342 (+7) Change $\sum_{i=1}^n$ to $\sum_{j=1}^n$

p. 351 (-1) Change $(\alpha\lambda)^k$ to $(\alpha\lambda)^t$

p. 353 (+9) Change $\phi(i_{t+1})$ to $\alpha\phi(i_{t+1})$

p. 354 (+5) Change $\phi(i_{k+1})$ to $\alpha\phi(i_{k+1})$

p. 369 (-16) Change Eq. (6.66) to read

$$\sum_{t=0}^{k} \phi(i_t)\tilde{q}(i_{t+1}, r_k) = \sum_{t \leq k,\, t \in T} \phi(i_t)c(i_{t+1}) + \left(\sum_{t \leq k,\, t \notin T} \phi(i_t)\phi(i_{t+1})' \right) r_k, \tag{6.66}$$

p. 370 (+13) Change $P(i_0 = i) > 0$ to $P(i_0 = i)$

p. 371 (+4) Change "... $q(1), \ldots, q(n)$." to "... $q(1), \ldots, q(n)$. We assume that $q_0(i)$ are chosen so that $q(i) > 0$ for all i [a stronger assumption is that $q_0(i) > 0$ for all i]."

p. 380 (-14) Change the first two sentences of the proof of Prop. 6.6.1(b) to read as follows:

(b) If $z \in \Re^n$ with $z \neq 0$ and $z \neq a\Pi Cz$ for all $a \geq 0$,

$$\|(1-\gamma)z + \gamma\Pi Cz\| < (1-\gamma)\|z\| + \gamma\|\Pi Cz\| \leq (1-\gamma)\|z\| + \gamma\|z\| = \|z\|, \tag{6.83}$$

where the strict inequality follows from the strict convexity of the norm, and the weak inequality follows from the non-expansiveness of ΠC. We also have $\|(1-\gamma)z+\gamma\Pi Cz\| < \|z\|$ if $z \neq 0$ and $z = a\Pi Cz$ for some $a \geq 0$, because then ΠH has a unique fixed point so $a \neq 1$, and ΠC is nonexpansive so $a < 1$.

p. 401 (+21) Along with Longstaff and Schwartz [LoS01], add reference to the following paper, which has similar content:

[TsV01] Tsitsiklis, J. N., and Van Roy, B., 2001. "Regression Methods for Pricing Complex American-Style Options," IEEE Trans. on Neural Networks, Vol. 12, pp. 694-703.

p. 404 (+1) Delete the 1st sentence: "Show ... $\Phi' V g/\Phi' V \Phi$."

p. 420 (+13) Change "$\overline{\mu}_k(u_k)$" to "$\overline{\mu}_k(x_k)$" (twice)

VOLUME 1 - 2ND EDITION

p. 10 (+21) Change "this true." to "this is true."

p. 109 (+4) The expression should read

$$g\big(x, \mu^*(t,x)\big) + \nabla_t J^*(t,x) + \nabla_x J^*\big(t,x\big)' f\big(x, \mu^*(t,x)\big),$$

p. 115 (-8) The equation should read

$$x^*(t) = x(0)e^{at} + \frac{b^2\xi}{2a}\big(e^{-at} - e^{at}\big),$$

p. 136, (+10) Change "the optimal $u^*(t)$" to "the sine of the slope of the optimal $x^*(t)$"

p. 150, (+10) Change "Eq. (4.12)" to "Eq. (4.11)"

p. 156, Fig. 4.2.1 Change $L(y)$ to $G_k(x_k)$

p. 157 (+13) Change $G(x_k)$ to $G_k(x_k)$

p. 160 (-11) Change

$$J_{N-1}(x) = \min\left[cx + G_{N-1}(x),\, \min_{y>x}\big[K + cy + G_{N-1}(y)\big]\right] - cx.$$

to

$$J_{N-1}(x) = \min\left[G_{N-1}(x),\, \min_{y>x}\big[K + G_{N-1}(y)\big]\right] - cx.$$

p. 168 The following is a cleaner version of the three paragraphs starting with the title "Asset Selling":

Asset Selling

As a first example, consider a person having an asset (say a piece of land) for which he is offered an amount of money from period to period. We assume that the offers, denoted $w_0, w_1, \ldots, w_{N-1}$, are random and independent, and take values within some bounded interval of noonnegative numbers ($w_k = 0$ could correspond to no offer received during the period). If the person accepts an offer, he can invest the money at a fixed rate of interest $r > 0$, and if he rejects the offer, he waits until the next period to consider the next offer. Offers rejected are not renewed, and we assume that the last offer w_{N-1} must be accepted if every prior offer has been rejected. The objective is to find a policy for accepting and rejecting offers that maximizes the revenue of the person at the Nth period.

The DP algorithm for this problem can be derived by elementary reasoning. As a modeling exercise, however, we will embed the problem in the framework of the basic problem by specifying the system and cost. We define the state space to be the real line, augmented with an additional state (call it T), which is a *termination state*. By writing that the system is at state $x_k = T$ at some time $k \leq N-1$, we mean that the asset has already been sold. By writing that the system is at a state $x_k \neq T$ at some time $k \leq N-1$, we mean that the asset has not been sold as yet and the offer under consideration is equal to x_k (and also equal to the kth offer w_{k-1}). We take $x_0 = 0$ (a fictitious "null" offer). The control space consists of two elements u^1 and u^2, corresponding to the decisions "sell" and "do not sell," respectively. We view w_k as the disturbance at time k.

With these conventions, we may write a system equation of the form

$$x_{k+1} = f_k(x_k, u_k, w_k), \qquad k = 0, 1, \ldots, N-1,$$

where the function f_k is defined via the relation

$$x_{k+1} = \begin{cases} T & \text{if } x_k = T, \text{ or if } x_k \neq T \text{ and } u_k = u^1 \text{ (sell)}, \\ w_k & \text{otherwise.} \end{cases}$$

Note that a sell decision at time k ($u_k = u^1$) accepts the offer w_{k-1}, and that no explicit sell decision is required to accept the last offer w_{N-1}, as it must be accepted by assumption if the asset has not yet been sold. The corresponding reward function may be written as

$$\mathop{E}_{\substack{w_k \\ k=0,1,\ldots,N-1}} \left\{ g_N(x_N) + \sum_{k=0}^{N-1} g_k(x_k, u_k, w_k) \right\}$$

where

$$g_N(x_N) = \begin{cases} x_N & \text{if } x_N \neq T, \\ 0 & \text{otherwise,} \end{cases}$$

$$g_k(x_k, u_k, w_k) = \begin{cases} (1+r)^{N-k} x_k & \text{if } x_k \neq T \text{ and } u_k = u^1 \text{ (sell)}, \\ 0 & \text{otherwise.} \end{cases}$$

p. 177 (+15) Change $dp(w)$ to $dP(w)$

p. 187 (+9) and p. 189 (+8) Change $N_k^{-1} C_k$ to N_k^{-1}

p. 217 (+1) Change "If involves" to "It involves"

p. 217 (-7) Change

$$\frac{P(x_1 = \overline{P}, G, G, S)}{P(G, G, S)}$$

to

$$\frac{P(x_1 = \overline{P}, G, G \mid S)}{P(G, G \mid S)}$$

p. 218 (+2) Change

$$\frac{P(x_1 = \overline{P}, G, B, S)}{P(G, B, S)}$$

to

$$\frac{P(x_1 = \overline{P}, G, B \mid S)}{P(G, B \mid S)}$$

p. 218 (+9) Change

$$\frac{P(x_1 = \overline{P}, G, G, C)}{P(G, G, C)}$$

to

$$\frac{P(x_1 = \overline{P}, G, G \mid C)}{P(G, G \mid C)}$$

p. 222 (-9) Change $x'_{N-1}K_{N-1}x_{N-1} \mid I_{N-2}\}$ to $E\{x'_{N-1}K_{N-1}x_{N-1} \mid I_{N-2}, u_{N-2}\}$

p. 244 (+9) The following is a cleaner version of the three pages that start with "**The Conditional State Distribution as a Sufficient Statistic**" title and end just before the "**The Conditional State Distribution Recursion**" title.

The Conditional State Distribution as a Sufficient Statistic

There are many different functions that can serve as sufficient statistics. The identity function $S_k(I_k) = I_k$ is certainly one of them. To obtain another important sufficient statistic, we assume that *the probability distribution of the observation disturbance* v_{k+1} *depends explicitly only on the immediately preceding state, control, and system disturbance* x_k, u_k, w_k, *and not on* $x_{k-1}, \ldots, x_0$, $u_{k-1}, \ldots, u_0$, $w_{k-1} \ldots, w_0$, $v_{k-1}, \ldots, v_0$. Under this assumption, it turns out that a sufficient statistic is given by the conditional probability distribution $P_{x_k|I_k}$ of the state x_k, given the information vector I_k. In particular, we will show that for all k and I_k, we have

$$J_k(I_k) = \min_{u_k \in U_k} H_k\big(P_{x_k|I_k}, u_k\big) = \overline{J}_k(P_{x_k|I_k}), \tag{5.34}$$

where H_k and $\overline{J}_k$ are appropriate functions.

To this end, we note an important fact that relates to state estimation of discrete-time stochastic systems: the conditional distribution $P_{x_k|I_k}$ can be generated recursively. In particular, it turns out that we can write for all k

$$P_{x_{k+1}|I_{k+1}} = \Phi_k\big(P_{x_k|I_k}, u_k, z_{k+1}\big),$$

where Φ_k is some function that can be determined from the data of the problem. Let us postpone a justification of this for the moment, and accept it for the purpose of the following discussion.

We note that to perform the minimization in Eq. (5.32), it is sufficient to know the distribution $P_{x_{N-1}|I_{N-1}}$ together with the distribution $P_{w_{N-1}|x_{N-1},u_{N-1}}$, which is part of the problem data. Thus, the minimization in the right-hand side of Eq. (5.32) is of the form

$$J_{N-1}(I_{N-1}) = \min_{u_{N-1}\in U_{N-1}} H_{N-1}\big(P_{x_{N-1}|I_{N-1}}, u_{N-1}\big) = \overline{J}_{N-1}(P_{x_{N-1}|I_{N-1}}),$$

for appropriate functions H_{N-1} and $\overline{J}_{N-1}$.

We now use induction, i.e., we assume that

$$J_{k+1}(I_{k+1}) = \min_{u_{k+1}\in U_{k+1}} H_{k+1}\big(P_{x_{k+1}|I_{k+1}}, u_{k+1}\big) = \overline{J}_{k+1}(P_{x_{k+1}|I_{k+1}}),$$

for appropriate functions H_{k+1} and $\overline{J}_{k+1}$, and we show that

$$J_k(I_k) = \min_{u_k\in U_k} H_k\big(P_{x_k|I_k}, u_k\big) = \overline{J}_k(P_{x_k|I_k}),$$

or appropriate functions H_k and $\overline{J}_k$.

Indeed, for a given I_k, the expression

$$\min_{u_k\in U_k} \mathop{E}_{x_k,w_k,z_{k+1}} \big\{g_k(x_k,u_k,w_k) + J_{k+1}(I_{k+1}) \mid I_k, u_k\big\}$$

in the DP equation (5.33) is written as

$$\min_{u_k\in U_k} \mathop{E}_{x_k,w_k,z_{k+1}} \big\{g_k(x_k,u_k,w_k) + J_{k+1}\big(\Phi_k\big(P_{x_k|I_k}, u_k, z_{k+1}\big)\big) \mid I_k, u_k\big\}.$$

In order to calculate the expression being minimized over u_k above, aside from $P_{x_k|I_k}$, we need the joint distribution

$$P(x_k, w_k, z_{k+1} \mid I_k, u_k)$$

or, equivalently,

$$P\big(x_k, w_k, h_{k+1}\big(f_k(x_k,u_k,w_k), u_k, v_{k+1}\big) \mid I_k, u_k\big).$$

By using Bayes' rule, this distribution can be expressed in terms of $P_{x_k|I_k}$, the given distributions

$$P(w_k \mid x_k, u_k), \qquad P\big(v_{k+1} \mid f_k(x_k,u_k,w_k), u_k, w_k\big),$$

and the system equation $x_{k+1} = f_k(x_k,u_k,w_k)$. Therefore the expression minimized over u_k can be written as a function of $P_{x_k|I_k}$ and u_k, and the DP equation (5.33) can be written as

$$J_k(I_k) = \min_{u_k\in U_k} H_k\big(P_{x_k|I_k}, u_k\big)$$

for a suitable function H_k. Thus the induction is complete and it follows that the distribution $P_{x_k|I_k}$ is a sufficient statistic.

Note that if the conditional distribution $P_{x_k|I_k}$ is uniquely determined by another expression $S_k(I_k)$, that is, $P_{x_k|I_k} = G_k\big(S_k(I_k)\big)$ for an appropriate function G_k, then $S_k(I_k)$ is also a sufficient statistic. Thus, for example, if we can show that $P_{x_k|I_k}$ is a Gaussian distribution, then the mean and the covariance matrix corresponding to $P_{x_k|I_k}$ form a sufficient statistic.

Regardless of its computational value, the representation of the optimal policy as a sequence of functions of the conditional probability distribution $P_{x_k|I_k}$,

$$\mu_k(I_k) = \overline{\mu}_k(P_{x_k|I_k}), \qquad k = 0, 1, \ldots, N-1,$$

is conceptually very useful. It provides a decomposition of the optimal controller in two parts:

(a) An *estimator*, which uses at time k the measurement z_k and the control u_{k-1} to generate the probability distribution $P_{x_k|I_k}$.

(b) An *actuator*, which generates a control input to the system as a function of the probability distribution $P_{x_k|I_k}$ (Fig. 5.4.1).

This interpretation has formed the basis for various suboptimal control schemes that separate the controller a priori into an estimator and an actuator and attempt to design each part in a manner that seems "reasonable." Schemes of this type will be discussed in Chapter 6.

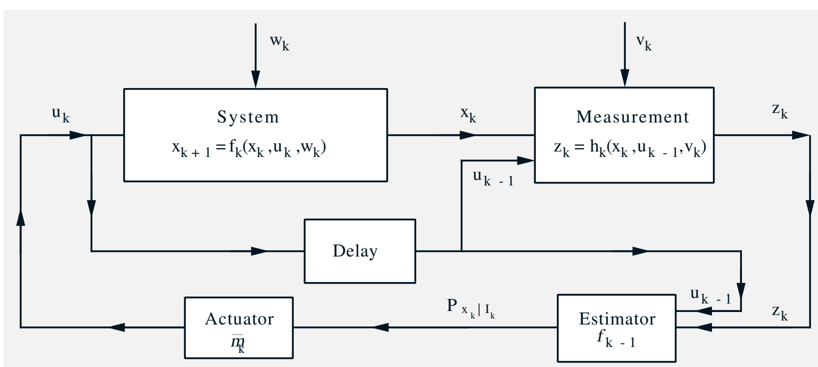

Figure 5.4.1 Conceptual separation of the optimal controller into an estimator and an actuator.

Alternative Perfect State Information Reduction

By using the sufficient statistic $P_{x_k|I_k}$ we can write the DP algorithm in an alternative form. Using Eq. (5.34), we have for $k < N - 1$

$$\overline{J}_k(P_{x_k|I_k}) = \min_{u_k \in U_k} \Big[\mathop{E}_{x_k, w_k, z_{k+1}} \big\{ g_k(x_k, u_k, w_k) + \overline{J}_{k+1}\big(\Phi_k(P_{x_k|I_k}, u_k, z_{k+1})\big) \mid I_k, u_k \big\} \Big]. \tag{5.35}$$

In the case where $k = N - 1$, we have

$$\overline{J}_{N-1}(P_{x_{N-1}|I_{N-1}}) = \min_{u_{N-1} \in U_{N-1}} \bigg[\mathop{E}_{x_{N-1}, w_{N-1}} \Big\{ g_N\big(f_{N-1}(x_{N-1}, u_{N-1}, w_{N-1})\big) + g_{N-1}(x_{N-1}, u_{N-1}, w_{N-1}) \mid I_{N-1}, u_{N-1} \Big\} \bigg]. \tag{5.36}$$

This DP algorithm yields the optimal cost as

$$J^* = \mathop{E}_{z_0}\big\{ \overline{J}_0(P_{x_0|z_0}) \big\},$$

where $\overline{J}_0$ is obtained by the last step, and the probability distribution of z_0 is obtained from the measurement equation $z_0 = h_0(x_0, v_0)$ and the statistics of x_0 and v_0.

By observing the form of Eq. (5.35), we note that it has the standard DP structure, except that $P_{x_k|I_k}$ plays the role of the "state." Indeed the role of the "system" is played by the recursive estimator of $P_{x_k|I_k}$,

$$P_{x_{k+1}|I_{k+1}} = \Phi_k\big(P_{x_k|I_k}, u_k, z_{k+1}\big),$$

and this system fits the framework of the basic problem (the role of control is played by u_k and the role of the disturbance is played by z_{k+1}). Furthermore, the controller can calculate (at least in principle) the state $P_{x_k|I_k}$ of this system at time k, so perfect state information prevails. Thus the alternate DP algorithm (5.34)-(5.35) may be viewed as the DP algorithm of the perfect state information problem that involves the above system, whose state is $P_{x_k|I_k}$, and an appropriately reformulated cost function. In the absence of perfect knowledge of the state, *the controller can be viewed as controlling the "probabilistic state" $P_{x_k|I_k}$ so as to minimize the expected cost-to-go conditioned on the information I_k available.*

p. 254 (-1) Change a_k to α_k

p. 255 (+6) Change $1 - \frac{1}{C}$ to $1 - \frac{I}{C}$

p. 261 (+8) Change "Exercise 5.6" to "Exercise 5.3"

p. 262 (+15) Change $y_0' K_k y_0$ to $y_0' K_0 y_0$

p. 266 (-5) Delete part (e) (it is correct only for open-loop policies)

p. 310 (11) Change "... tentative backed-up score $TBS(n)$ of position n to ∞ if it is White's turn to move at n and to $-\infty$ if it is Black's turn ..." to "...tentative backed-up score $TBS(n)$ of position n to $-\infty$ if it is White's turn to move at n and to ∞ if it is Black's turn ..."

p. 316 (+16) Change "of question" to "of questions"

p. 349 (+15) Change "Section 1.3" to "Section 2.3"

p. 353 (+12) Change "CEC is 1." to "CEC with nominal values $\overline{w}_0 = \overline{w}_1 = 0$ is 1."

p. 357 (+13) Replace Exercise 6.13 by the following:

6.13 (Discretization of Convex Problems)

Consider a problem with state space S, for all k, where S is a convex subset of $\Re^n$. Suppose that $\hat{S} = \{y_1, \ldots, y_M\}$ is a finite subset of S such that S is the convex hull of $\hat{S}$, and consider a one-step lookahead policy based on approximated cost-to-go functions $\tilde{J}_0, \tilde{J}_1, \ldots, \tilde{J}_N$ defined as follows:

$$\tilde{J}_N(x) = g_N(x), \qquad \forall\ x \in S,$$

and for $k = 1, \ldots, N-1$,

$$\tilde{J}_k(x) = \min \left\{ \sum_{i=1}^{M} \lambda_i \hat{J}_k(y_i) \;\middle|\; \sum_{i=1}^{M} \lambda_i y_i = x,\ \sum_{i=1}^{M} \lambda_i = 1,\ \lambda_i \geq 0,\ i = 1, \ldots, M \right\},$$

where $\hat{J}_k(x)$ is defined by

$$\hat{J}_k(x) = \min_{u \in U_k(x)} E\Big\{ g_k(x, u, w_k) + \tilde{J}_{k+1}\big(f_k(x, u, w_k)\big) \Big\}, \qquad \forall\ i = 1, \ldots, M.$$

Thus $\tilde{J}_k$ is obtained from $\tilde{J}_{k+1}$ as a "grid-based" convex piecewise linear approximation to $\hat{J}_k$ based on the M values

$$\hat{J}_k(y_1), \ldots, \hat{J}_k(y_M).$$

Assume that the cost functions g_k and the system functions f_k are such that the function $\hat{J}_k$ is real-valued and convex over S whenever $\tilde{J}_{k+1}$ is real-valued and convex over S. Show that the cost-to-go functions $\overline{J}_k(x_k)$ corresponding to the one-step lookahead policy satisfies for all $x \in S$

$$\overline{J}_k(x) \leq \hat{J}_k(x) \leq \tilde{J}_k(x), \qquad k = 0, 1, \ldots, N-1.$$

Hint: Use Prop. 6.3.1.

p. 373 (-1) Replace $N_*(j)$ by $N^*(j)$

p. 374 (+3), (+8) Replace $N_*(j)$ by $N^*(j)$

p. 396 (+6) The expression should read

$$\overline{\tau}_i(u) = \sum_{j=1}^{n} \int_0^{\infty} \tau dQ_{ij}(\tau, u),$$

p. 402 (+15), (+19), (-5) Change $g(i,u)$ to $G(i,u)$

p. 402 (-5) After Eq. (7.54), add the following sentence: If there is an "instantaneous" one-stage cost $\hat{g}(i,u)$, the term $G(i,u)$ should be replaced by $\hat{g}(i,u) + G(i,u)$ in this equation.

p. 408 (+2) Change (7.6) to (7.17)

p. 451 (+12) Change $C_k N_k^{-1}$ to $C_k' N_k^{-1}$

p. 476 (-11) Change

$$P_{d_1} \preceq P_{d_2} \ \text{ if and only if } \ E\big\{U\big(f(d_1,n)\big) \mid d_1\big\} \le \big\{U\big(f(d_2,n)\big) \mid d_2\big\}.$$

to

$$P_{d_1} \preceq P_{d_2} \ \text{ if and only if } \ E\big\{U\big(f(d_1,n)\big) \mid d_1\big\} \le E\big\{U\big(f(d_2,n)\big) \mid d_2\big\}.$$

VOLUME 2 - 2ND EDITION

p. 7 (+19) Change "operation D can be performed only after operation B has been performed" to "operation D can be performed only after operation C has been performed"

p. 50 (+17) Change "Tsitsiklis" to "Castanon"

p. 64 (+13) Change the first four lines of the proof as follows:

Proof: In view of Eqs. (1.59) and (1.65), existence of a PPR policy is equivalent to having, for all i,

$$\max\left\{M, \max_{j\neq i} L^j(x, M, J)\right\} \geq L^i(x, M, J), \qquad \text{for all } x \text{ with } x^i \in S^i, \tag{1.66}$$

$$M \leq L^i(x, M, J), \qquad \text{for all } x \text{ with } x^i \notin S^i, \tag{1.67}$$

p. 70 (+10) Change p_i to P_i

p. 74 (+1) Change " ... , [VeP84], and Verd'u and Poor [VeP87]." to " ... , and Verd'u and Poor [VeP84], [VeP87]."

p. 80 (+17,+19,27) Change 1 to s

p. 84 (+10) Change the equation to

$$T_{\overline{\mu}} J = TJ. \tag{1.87}$$

p. 94 (-1) Change equation to

$$J_\mu(1) = -(1-u^2)u + (1-u^2)J_\mu(1)$$

p. 95 (+2) Change equation to

$$J_\mu(1) = -\frac{1-u^2}{u}.$$

p. 123 (+6) Change "thrtwe" to (2.33)

p. 136 (+3) Change "an optimal proper policy" to "a proper policy"

p. 136 (+11) Change "optimal and proper." to "optimal."

p. 136 (-10) Change the last four lines of the hint to:

By taking limit superior as $N \to \infty$, we obtain $J^* \geq TJ^* \geq J_\mu$. Therefore, μ is an optimal policy and we have $J^* = TJ^*$. For the rest of the proof follow the line of proof of Prop. 2.1.2.

p. 181 (-2) Change "discount factor" to "discount factor with $\alpha < 2$"

p. 190 (-1) Change "no optimal policy (stationary or not)." to "no optimal stationary policy."

p. 206 (-1) Change "N_1" to "$N - 1$"

p. 208 (-12) Change Exercise number to 4.32.

p. 234 (+20) Change "Prop. 4.3.3" to "Prop. 4.2.6"

p. 235 (+14) Change "the probabilities $q(i, u)$, $u \in U(i)$." to "some probabilities."

p. 247 (-4) Change "for some $\beta > 0$" to "for some $\beta \in (0, 1)$"

p. 264 (-6) Change a_k to α_k

p. 265 (+3) Change $1 - \frac{1}{C}$ to $1 - \frac{I}{C}$