Dynamic Programming Deconstructed

Qingyin Ma and John Stachurski

July 2018

Introduction

Consider a worker who observes wage offer w_t and either

- **rejects**, gets unemployment compensation η_t, waits for next period
- **accepts**, works forever at current offer w_t

Bellman's equation, assuming IID shocks:

$$v(w, \eta) = \max\left\{\frac{u(w)}{1-\beta}, \, u(\eta) + \beta \mathbb{E}\, v(w', \eta')\right\}$$

We can simplify this problem by a "time shift"

Examples. Start the clock when

- agent has received current offer and rejected it
- agent has received current offer, rejected it and received unemployment compensation

Leads to alternative "versions" of the Bellman equation

Benefit: we have **multiple angles of attack** on the same problem

In fact economists have come up with many tricks to simplify Bellman equations in different applications

Collectively, we call them "plan factorizations"

The basic idea:

- rearrange the Bellman equation into a more advantageous form. . .
- while preserving its link to optimality

Examples from a variety of applications:

- Jovanovic (1982)
- Rust (1987)
- Aguirregabiria and Mira (2002)

$\vdots \leftarrow$ insert many names

- Fajgelbaum, Schaal and Taschereau-Dumouchel (2017)
- Winberry (2018)
- Kristensen, Mogensen, Moon and Schjerning (2018)
- etc., etc.

We compose a theoretical framework for plan decomposition that

- contains all of the existing manipulations as special cases
- extends to nonseparable preferences
  - recursive preferences
  - risk-sensitive control
  - desire for robustness, ambiguity, etc.
- provides a full set of optimality results
- obtains relative rates of convergence
- obtains new results by exploiting plan factorizations in novel ways

We compose a theoretical framework for plan decomposition that

- contains all of the existing manipulations as special cases
- extends to nonseparable preferences
    - recursive preferences
    - risk-sensitive control
    - desire for robustness, ambiguity, etc.
- provides a full set of optimality results
- obtains relative rates of convergence
- obtains new results by exploiting plan factorizations in novel ways

We compose a theoretical framework for plan decomposition that

- contains all of the existing manipulations as special cases
- extends to nonseparable preferences
  - recursive preferences
  - risk-sensitive control
  - desire for robustness, ambiguity, etc.
- provides a full set of optimality results
- obtains relative rates of convergence
- obtains new results by exploiting plan factorizations in novel ways

We compose a theoretical framework for plan decomposition that

- contains all of the existing manipulations as special cases
- extends to nonseparable preferences
    - recursive preferences
    - risk-sensitive control
    - desire for robustness, ambiguity, etc.
- provides a full set of optimality results
- obtains relative rates of convergence
- obtains new results by exploiting plan factorizations in novel ways

We compose a theoretical framework for plan decomposition that

- contains all of the existing manipulations as special cases
- extends to nonseparable preferences
    - recursive preferences
    - risk-sensitive control
    - desire for robustness, ambiguity, etc.
- provides a full set of optimality results
- obtains relative rates of convergence
- obtains new results by exploiting plan factorizations in novel ways

Examples from the Literature

Recall the Bellman equation of the McCall model discussed above:

$$v(w, \eta) = \max \left\{ \frac{u(w)}{1-\beta}, \, u(\eta) + \beta \mathbb{E}\, v(w', \eta') \right\}$$

A functional equation in **two** dimensions...

Now let

$$g(\eta) := u(\eta) + \beta \mathbb{E}\, v(w', \eta') \tag{1}$$

Then Bellman's equation becomes

$$v(w, \eta) = \max \left\{ \frac{u(w)}{1 - \beta}, g(\eta) \right\} \tag{2}$$

Combining (1) and (2) to eliminate v gives

$$g(\eta) = u(\eta) + \beta \mathbb{E} \max \left\{ \frac{u(w')}{1 - \beta}, g(\eta') \right\}$$

A functional equation in **one** dimension

We can do still better using Rust's "expected value" approach

Let $h := \beta \mathbb{E}\, v(w', \eta')$, so that

$$v(w, \eta) = \max \left\{ \frac{u(w)}{1-\beta}, \, u(\eta) + h \right\}$$

Now eliminate v to get

$$h = \beta \mathbb{E} \, \max \left\{ \frac{u(w')}{1-\beta}, \, u(\eta') + h \right\}$$

A functional equation in **zero** dimensions

Example. (Optimal savings) Consider

$$v(w, z, \eta) = \max_{c, w', \ell} \left\{ u(c, \ell) + \beta \, \mathbb{E}_z \, v(w', z', \eta') \right\}$$

subject to

$$w' + c \leqslant Rw + q(z, \eta)\ell, \quad c, w' \geqslant 0 \quad \text{and} \quad 0 \leqslant \ell \leqslant 1.$$

Here

- $\{\eta_t\}$ are IID innovations and $\{z_t\}$ is Markov
- $\mathbb{E}_z$ conditions on current state z

The BE is a three dimensional functional equation

Analogous to Winberry (2018), let

$$h(w,z) := \mathbb{E}_z \, v(w,z,\eta)$$

Use the law of iterated expectations to write the BE as

$$v(w,z,\eta) = \max_{c,w',\ell} \left\{ u(c,\ell) + \beta \, \mathbb{E}_z \, \mathbb{E}_{z'} \, v(w',z',\eta') \right\}$$

and rearrange to get

$$h(w,z) = \mathbb{E}_z \max_{c,w',\ell} \left\{ u(c,\ell) + \beta \mathbb{E}_z \, h(w',z') \right\}$$

subject to

$$w' + c \leqslant Rw + q(z,\eta)\ell, \quad c, w' \geqslant 0 \quad \text{and} \quad 0 \leqslant \ell \leqslant 1$$

A two dimensional functional equation. . .

Questions!

- At what point does the standard theory break down?
- Does contractivity imply optimality?
- Does iteration converge at the same rate?
- Might one sequence converge while another does not?
- Do the answers to these questions change if we have
  - EZ preferences?
  - ambiguity?
  - desire for robustness?
  - risk-sensitive preferences?

Questions!

- At what point does the standard theory break down?
- Does contractivity imply optimality?
- Does iteration converge at the same rate?
- Might one sequence converge while another does not?
- Do the answers to these questions change if we have
  - EZ preferences?
  - ambiguity?
  - desire for robustness?
  - risk-sensitive preferences?

Questions!

- At what point does the standard theory break down?
- Does contractivity imply optimality?
- Does iteration converge at the same rate?
- Might one sequence converge while another does not?
- Do the answers to these questions change if we have
  - EZ preferences?
  - ambiguity?
  - desire for robustness?
  - risk-sensitive preferences?

Questions!

- At what point does the standard theory break down?
- Does contractivity imply optimality?
- Does iteration converge at the same rate?
- Might one sequence converge while another does not?
- Do the answers to these questions change if we have
  - EZ preferences?
  - ambiguity?
  - desire for robustness?
  - risk-sensitive preferences?

Questions!

- At what point does the standard theory break down?
- Does contractivity imply optimality?
- Does iteration converge at the same rate?
- Might one sequence converge while another does not?
- Do the answers to these questions change if we have
    - EZ preferences?
    - ambiguity?
    - desire for robustness?
    - risk-sensitive preferences?

New Applications of Plan Factorizations

Plan factorizations have been used in the past to

- reduce dimension
- obtain more smoothness
- simplify estimation

Can they help us in **other** ways???

Example. Benhabib et al. (JET, 2015) study a Bewley model with financial income risk and Bellman equation

$$v(w) = \max_{0 \leqslant s \leqslant w} \left\{ u(w - s) + \beta \, \mathbb{E} \, v(R's + y') \right\}$$

Here

- $\{R_t\}$ and $\{y_t\}$ are nonnegative and IID
- CRRA utility $u(c) = c^{1-\gamma}/(1-\gamma)$ with $\gamma > 1$

One difficulty: v unbounded below, so can't use

- standard DP arguments with sup norms...
- or even weighted sup norm methods

Can a plan factorization help?

Let

$$g(s) := \beta \mathbb{E}\, v(R's + y')$$

so that

$$v(w) = \max_{0 \leqslant s \leqslant w} \{u(w - s) + g(s)\}$$

Manipulating to eliminate v gives

$$g(s) = \beta\, \mathbb{E} \max_{0 \leqslant s' \leqslant R's + y'} \left\{ u(R's + y' - s') + g(s') \right\}$$

Corresponding "Bellman-like" operator is

$$Sg(s) = \beta\, \mathbb{E} \max_{0 \leqslant s' \leqslant R's + y'} \left\{ u(R's + y' - s') + g(s') \right\}$$

If $g(s) \geqslant M$, then

$$Sg(s) = \beta \, \mathbb{E} \max_{0 \leqslant s' \leqslant R's + y'} \left\{ u(R's + y' - s') + g(s') \right\}$$

$$\geqslant \beta \, \mathbb{E} \max_{0 \leqslant s' \leqslant R's + y'} \left\{ u(R's + y' - s') + M \right\}$$

$$= \beta \, \mathbb{E} \left\{ u(R's + y') + M \right\}$$

$$\therefore \quad Sg(s) \geqslant \beta \, \mathbb{E} \, u(y') + M$$

Thus, $\mathbb{E} \, u(y') > -\infty \implies$

- Sg also bounded below!
- and a sup norm contraction...

Questions!

- S is a contraction but what does this imply?
- If one Bellman-like equation has a unique fixed point, do all?
- Does this lead us to optimality?
  - After all, the original Bellman op. is not a sup norm contraction
- Is that still true if we add in, say, risk sensitivity?
- Are there more plan factorizations we should try?
- Can they help us in other ways?

Questions!

- S is a contraction but what does this imply?
- If one Bellman-like equation has a unique fixed point, do all?
- Does this lead us to optimality?
  - After all, the original Bellman op. is not a sup norm contraction
- Is that still true if we add in, say, risk sensitivity?
- Are there more plan factorizations we should try?
- Can they help us in other ways?

Questions!

- S is a contraction but what does this imply?
- If one Bellman-like equation has a unique fixed point, do all?
- Does this lead us to optimality?
  - After all, the original Bellman op. is not a sup norm contraction
- Is that still true if we add in, say, risk sensitivity?
- Are there more plan factorizations we should try?
- Can they help us in other ways?

Questions!

- S is a contraction but what does this imply?
- If one Bellman-like equation has a unique fixed point, do all?
- Does this lead us to optimality?
  - After all, the original Bellman op. is not a sup norm contraction
- Is that still true if we add in, say, risk sensitivity?
- Are there more plan factorizations we should try?
- Can they help us in other ways?

Questions!

- S is a contraction but what does this imply?
- If one Bellman-like equation has a unique fixed point, do all?
- Does this lead us to optimality?
  - After all, the original Bellman op. is not a sup norm contraction
- Is that still true if we add in, say, risk sensitivity?
- Are there more plan factorizations we should try?
- Can they help us in other ways?

Questions!

- S is a contraction but what does this imply?
- If one Bellman-like equation has a unique fixed point, do all?
- Does this lead us to optimality?
  - After all, the original Bellman op. is not a sup norm contraction
- Is that still true if we add in, say, risk sensitivity?
- Are there more plan factorizations we should try?
- Can they help us in other ways?

Theory

In what follows, a dynamic programming problem consists of

- a **state space** $\mathbb{X}$
- an **action space** $\mathbb{A}$
- a **feasible correspondence** Γ
- a set of **candidate value functions** $\mathbb{V}$
- a **state-action aggregator** Q

The Bellman equation:

$$v(x) = \max_{a \in \Gamma(x)} Q(x, a, v) \quad \text{for all } \; x \in \mathbb{X}$$

The **Bellman operator** is

$$T\,v(x) = \max_{a \in \Gamma(x)} Q(x, a, v)$$

A v-**greedy** policy is a feasible policy σ such that, for all x,

$$\sigma(x) \in \operatorname*{argmax}_{a \in \Gamma(x)} Q(x, a, v)$$

- Treats v as the value function and makes best choice

Can accommodate

- recursive preferences
- risk-sensitive preferences
- robustness
- ambiguity
- etc.

Just modify the definition of Q...

Example. Benhabib et al. (JET, 2015) continued:

- $x = w$
- $a = s$
- $\Gamma(w) = [0, w]$
- $Q(w, s, v) = u(w - s) + \beta \, \mathbb{E} \; v(R's + y')$

Example. Benhabib et al. with CES aggregator:

Same, except

$$Q(w, s, v) = \left\{ (1 - \beta) u(w - s)^{\rho} + \beta \left[\mathbb{E} \; v(R's + y') \right]^{\rho} \right\}^{1/\rho}$$

We define a **plan factorization** to be a pair of operators (W_0, W_1) be such that

$$(W_1 \, W_0 \, v)(x, a) = Q(x, a, v), \quad \forall \text{ feasible } (x, a) \text{ and } \forall v \in \mathbb{V}$$

- $W_0 \circ W_1$ represents future value conditional on current action
- An arbitrary decomposition

Letting

$$Mh(x) := \max_{a \in \Gamma(x)} h(x, a)$$

the Bellman operator can be written as

$$T = M \circ W_1 \circ W_0$$

Example. Recall that in Benhabib et al. (JET, 2015),

$$Q(w,s,v) = u(w-s) + \beta \, \mathbb{E} \, v(R's + y')$$

Our plan factorization is (W_0, W_1) where

- $(W_0 v)(s) = g(s) := \beta \, \mathbb{E} \, v(R's + y')$
- $(W_1 g)(w,s) := u(w-s) + g(s)$

As required,

$$(W_1 \, W_0 \, v)(w,s) = Q(w,s,v)$$

Rearranging the Bellman Operator

Recall that

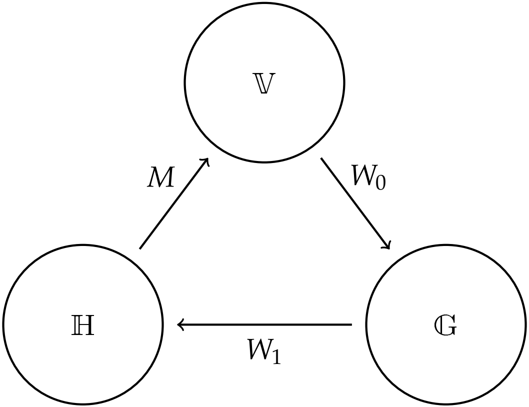
$$T = M \circ W_1 \circ W_0 \text{ on } \mathbb{V}$$

Now consider

$$S = W_0 \circ M \circ W_1 \text{ on } \mathbb{G}$$

Here $\mathbb{G} := W_0 \mathbb{V}$

We call S the **refactored Bellman operator**



$\mathbb{V}$
M
W_0
$\mathbb{H}$
$\mathbb{G}$
W_1

Example. Recall again Benhabib et al. (JET, 2015), with "Bellman-like" operator

$$Sg(s) = \beta \, \mathbb{E} \max_{0 \leqslant s' \leqslant R's + y'} \left\{ u(R's + y' - s') + g(s') \right\} \qquad (3)$$

Our plan factorization was

- $(W_0 v)(s) = g(s) := \beta \, \mathbb{E} \; v(R's + y')$
- $(W_1 g)(w, s) := u(w - s) + g(s)$

It follows that S in (3) can be expressed as the factorized Bellman operator

$$S = W_0 \circ M \circ W_1 \text{ on } \mathbb{G}$$

For each $\sigma \in \Sigma :=$ **feasible policies**, define

$$M_\sigma v(x) := v(x, \sigma(x))$$

Policy operators

$$T_\sigma = M_\sigma \circ W_1 \circ W_0 \text{ on } \mathbb{V}$$

$$S_\sigma = W_0 \circ M_\sigma \circ W_1 \text{ on } \mathbb{G}$$

Let

- $v_\sigma :=$ fixed point of T_σ and $v^* := \sup_\sigma v_\sigma =$ **value function**
- $g_\sigma :=$ fixed point of S_σ
- $g^* := \sup_\sigma g_\sigma =$ **refactored value function**

Question: Does the refactored Bellman operator have "equal rights"?

For example, suppose we obtain contraction results for S

Does this mean we can

- proceed as we would for T and
- be confident that the result is approximately optimal?

Answer: Yes, if W_0 is isotone on $\mathbb{V}$ and W_1 is isotone on $\mathbb{G}$

Let's call this **joint isotonicity**

Main Results

Thm. Under joint isotonicity, the next two statements are equivalent:

1. v^* satisfies the Bellman equation
2. g^* satisfies refactored Bellman equation

In particular, if g^* satisfies refactored Bellman equation, then

- Bellman's principle of optimality holds
- set of optimal policies is nonempty

So when does $Sg^* = g^*$?

Main Results

Thm. Under joint isotonicity, the next two statements are equivalent:

1. v^* satisfies the Bellman equation
2. g^* satisfies refactored Bellman equation

In particular, if g^* satisfies refactored Bellman equation, then

- Bellman's principle of optimality holds
- set of optimal policies is nonempty

So when does $Sg^* = g^*$?

Thm. Assume joint isotonicity. If $\exists$ a topology τ on $\mathbb{G}$ under which

- the pointwise partial order is closed and
- S and $\{S_\sigma\}_{\sigma \in \Sigma}$ are asymptotically stable on $\mathbb{G}$,

then g^* satisfies the refactored Bellman equation

Moreover,

1. A feasible policy is optimal if and only if it is g^*-greedy
2. $S^k g \to g^*$ as $k \to \infty$ for all $g \in \mathbb{G}$
3. At least one optimal policy exists

Translation: Under joint montonicity, refactored version has equal rights

You can

- pick the pick the factorization you want to work with
- establish contraction (or similar) results in that setting
- proceed in the obvious way

We also prove that iterates are “conjugate”

Translation: Freely use the factorization that produces the most efficient Bellman operator

Joint isotonicity not required for above

Plus other results on algorithms, optimality. . . see paper

Application

Bäuerle and Jaśkiewicz (JET, 2018) study a growth model with **risk-sensitive preferences**

The Bellman equation is

$$v(x) = \sup_{y \in [0,x]} \left\{ u(x-y) - \frac{\beta}{\gamma} \log \int \exp\left(-\gamma\, v[f(y,z)]\right) \mu(\mathrm{d}z) \right\}$$

They provide conditions for the validity of

- value function iteration
- Bellman's principle of optimality

One short-coming: utility must be bounded below

We generalize their results to include utility unbounded below

Method:

1. Implement a plan factorization to obtain the refactored Bellman equation

$$g(y) = -\frac{1}{\gamma} \times$$

$$\log \int \exp \left\{ -\gamma \sup_{y' \in [0, f(y,z)]} \left[u \left(f(y,z) - y' \right) + \beta g(y') \right] \right\} \mu(\mathrm{d}z)$$

2. Establish the conditions of the preceding theorem

In doing so we exploit the fact that the solution g is naturally bounded below