Back Substitution to Solve Recurrence Relations

Let's say you wanted a closed-form solution (explicit formula) for the recurrence relation:

$$\begin{aligned} R_1 &= 1 \\ R_n &= 2R_{n-1} + 2 \end{aligned}$$

How could we go about it? One method is known as back-substitution or sometimes substitute-and-simplify. In this method we start with the general form of the recurrence (R_n here) and substitute for the prior terms from the right side what they would be equal to:

$$\begin{aligned} R_n &= 2R_{n-1} + 2 \\ &= 2(2R_{n-2} + 2) + 2 \end{aligned}$$

At first glance, it appears we've just made it messier, but we then proceed to simplify it somewhat:

$$\begin{aligned} R_n &= 2R_{n-1} + 2 \\ &= 2(2R_{n-2} + 2) + 2 \\ &= 2^2 R_{n-2} + 2^2 + 2 \end{aligned}$$

The trick here is to never oversimplify things. The usual urge amongst young practitioners is to simplify it to the max like so: $4R_{n-2} + 6 = 2(R_{n-2} + 3)$. This oversimplification will lead us to nowhere fast. What we are looking for is a helpful pattern to help make this into that closed or explicit form. Let's just follow along for a bit:

$$\begin{aligned} R_n &= 2R_{n-1} + 2 & (*) \\ &= 2(2R_{n-2} + 2) + 2 \\ &= 2^2 R_{n-2} + 2^2 + 2 & (*) \\ &= 2^2(2R_{n-3} + 2) + 2^2 + 2 \\ &= 2^3 R_{n-3} + 2^3 + 2^2 + 2 & (*) \end{aligned}$$

Already from the starred lines we see a pattern forming. Whenever the subscript is having k subtracted, we are multiplying the R_{n-k} term by 2^k and adding a new 2^k term to the end. That is, we see:

$$
\begin{aligned}
R_n &= 2R_{n-1} + 2 \\
&= 2(2R_{n-2} + 2) + 2 \\
&= 2^2 R_{n-2} + 2^2 + 2 \\
&= 2^2(2R_{n-3} + 2) + 2^2 + 2 \\
&= 2^3 R_{n-3} + 2^3 + 2^2 + 2 \\
&\vdots \\
&= 2^k R_{n-k} + 2^k + \cdots + 2^2 + 2 \qquad (*)
\end{aligned}
$$

In the new starred line — the extrapolated pattern line — we have the general form of the substitutions. But where does this lead us ultimately? We need to see how far k can take us. Since this recurrence started with $n = 1$ as the base case, we can take the k value all the way to $n - 1$ before $n - k$ becomes undefined. That is, when we substitute in $k = n - 1$ we get $n - (n - 1) = 1$ as the subscript which is the initial value of the recurrence:

$$
\begin{aligned}
R_n &= 2R_{n-1} + 2 \\
&= 2(2R_{n-2} + 2) + 2 \\
&= 2^2 R_{n-2} + 2^2 + 2 \\
&= 2^2(2R_{n-3} + 2) + 2^2 + 2 \\
&= 2^3 R_{n-3} + 2^3 + 2^2 + 2 \\
&\vdots \\
&= 2^k R_{n-k} + 2^k + \cdots + 2^2 + 2 \\
&\vdots \\
&= 2^{n-1} R_1 + 2^{n-1} + \cdots + 2^2 + 2 \\
&= 2^{n-1} R_1 + \sum_{j=1}^{n-1} 2^j
\end{aligned}
$$

The last line changes the series of powers of 2 into a nice partial sum of a geometric sequence. It is just missing the $2^0 = 1$ term when $j = 0$. We can fix this by both adding and subtracting 1 from the right side and including the added one inside the partial sum as a 2^0 term. Let's also substitute in $R_1 = 1$ to get rid of that as well:

$$
\begin{aligned}
R_n &= 2R_{n-1} + 2 \\
&= 2(2R_{n-2} + 2) + 2 \\
&= 2^2 R_{n-2} + 2^2 + 2 \\
&= 2^2(2R_{n-3} + 2) + 2^2 + 2 \\
&= 2^3 R_{n-3} + 2^3 + 2^2 + 2 \\
&\vdots \\
&= 2^k R_{n-k} + 2^k + \cdots + 2^2 + 2 \\
&\vdots \\
&= 2^{n-1} R_1 + 2^{n-1} + \cdots + 2^2 + 2 \\
&= 2^{n-1} R_1 + \sum_{j=1}^{n-1} 2^j \\
&= 2^{n-1} + \sum_{j=0}^{n-1} 2^j - 1
\end{aligned}
$$

Now we can simplify that partial sum with our handy formula from the summation section:

$$
\begin{aligned}
R_n &= 2R_{n-1} + 2 \\
&= 2(2R_{n-2} + 2) + 2 \\
&= 2^2 R_{n-2} + 2^2 + 2 \\
&= 2^2(2R_{n-3} + 2) + 2^2 + 2 \\
&= 2^3 R_{n-3} + 2^3 + 2^2 + 2 \\
&\vdots \\
&= 2^k R_{n-k} + 2^k + \cdots + 2^2 + 2 \\
&\vdots \\
&= 2^{n-1} R_1 + 2^{n-1} + \cdots + 2^2 + 2 \\
&= 2^{n-1} R_1 + \sum_{j=1}^{n-1} 2^j \\
&= 2^{n-1} + \sum_{j=0}^{n-1} 2^j - 1 \\
&= 2^{n-1} + \frac{2^n - 1}{2 - 1} - 1 \\
&= 2^n + 2^{n-1} - 2 = 2^{n-1}(2+1) - 2 = 3 \cdot 2^{n-1} - 2
\end{aligned}
$$

Now, all this pattern finding is really just a conjecture on our part so we should prove this is so by induction, of course. Let's have a go, shall we?

Proof:

Base Case: $n = 1$
Left side: $R_1 = 1$
Right side: $3 \cdot 2^{1-1} - 2 = 3 \cdot 2^0 - 2 = 3 - 2 = 1$
These are equal, so the base case is proved.

Induction Step: Assume we know that for some $k \geq 1$ that $R_k = 3 \cdot 2^{k-1} - 2$. We will show that $R_{k+1} = 3 \cdot 2^{k+1-1} - 2 = 3 \cdot 2^k - 2$.

$$\begin{aligned}
R_{k+1} &= 2R_k + 2 && \text{back to recurrence definition} \\
&= 2(3 \cdot 2^{k-1} - 2) + 2 && \text{by inductive hypothesis} \\
&= 3 \cdot 2 \cdot 2^{k-1} - 4 + 2 \\
&= 3 \cdot 2^{k-1+1} - 2 \\
&= 3 \cdot 2^k - 2
\end{aligned}$$

Thus we have proved our conjecture and the closed form (or explicit formula) for R_n is shown to be $3 \cdot 2^{n-1} - 2$.

Induction proof aside, hopefully the substitute-and-simplify method made sense and you can use it to solve simple cases and maybe even not-so-simple cases alike.

Exercises

Solve each of the following recurrence relations by back-substitution. (Don't forget to prove your result by induction afterward!)

1. $a_0 = 4$ and for $n \geq 1$, $a_n = 2a_{n-1} + 1$
2. $b_0 = 2$ and for $n \geq 1$, $b_n = n \cdot b_{n-1}$ $\quad *$
3. $s_0 = 3$ and for $n \geq 1$, $s_n = s_{n-1} + n$
4. $t_0 = 0$ and $t_1 = 1$ and for $n \geq 2$, $t_n = t_{n-2} - 2$

Solutions to Starred Exercises

2. Solving for b_n:

$$\begin{aligned}
b_n &= n \cdot b_{n-1} \\
&= n((n-1)b_{n-2}) \\
&= n(n-1)b_{n-2} \\
&= n(n-1)((n-2)b_{n-3}) \\
&= n(n-1)(n-2)b_{n-3} \\
&\vdots \\
&= n(n-1)(n-2)\cdots(n-k+1)b_{n-k} \\
&\vdots \\
&= n(n-1)(n-2)\cdots 1 b_0 \\
&= 2n!
\end{aligned}$$

Proof:

Base Case: $n = 0$
Left side: $b_0 = 2$
Right side: $2 \cdot 0! = 2 \cdot 1 = 2$
These are equal, so the base case is proved.

Induction Step: Assume we know that for some $k \geq 0$ that $b_k = 2k!$. We will show that $b_{k+1} = 2(k+1)!$.

$$\begin{aligned}
b_{k+1} &= (k+1) \cdot b_k && \text{by recurrence definition} \\
&= (k+1) \cdot (2k!) && \text{by inductive hypothesis} \\
&= 2(k+1)k! \\
&= 2(k+1)!
\end{aligned}$$

Thus we have proved our conjecture and the closed form for $b_n = 2n!$.