23 Lecture #16: Wednesday, April 8th, 2026

23.1 The optimal in the relaxation procedure

We now ask whether there is a way to determine an optimal λ in the relaxation procedure (see Definition 22.13). We begin with the following observation. Let us return to our fixed-point problem $\varphi(x) = x$. Since we can rewrite this equation as $\varphi(x) - x = 0$, we may look for a point x_0 such that $x_0 > \varphi(x_0)$ and another point x_1 such that $x_1 < \varphi(x_1)$. This implies that $\varphi(x_0) - x_0 < 0$ and $\varphi(x_1) - x_1 > 0$. By the Intermediate Value Theorem, there exists $x_f \in [x_0, x_1]$ such that $\varphi(x_f) - x_f = 0$, that is, x_f is a fixed point of φ. Therefore, we can narrow down the interval to which x_f belongs.

This implies that we can start our iterations from a point close to the unknown fixed point x_f, say x_0. We may then ask whether the derivative of φ is small near this point x_0. If it is, then there is a chance that the iteration $x_{k+1} = \varphi(x_k)$ will converge. If this is not the case, one should instead consider using relaxation. Thus, we define

$$\varphi_\lambda(x) = \lambda\varphi(x) + (1 - \lambda)x.$$

Since we are interested in flat functions (see Theorem 22.11), our goal is to choose λ so that the function φ_λ is as flat as possible. This occurs when $\varphi_\lambda' = 0$. The key point, however, is to determine around which point this should happen. In the best-case scenario, this would occur at x_f, but since this point is unknown, we instead consider a nearby point x_0 and study the equation $\varphi_\lambda'(x_0) = 0$. In this case, we obtain

$$\lambda\varphi'(x_0) + (1 - \lambda) \cdot 1 = 0 \quad \Rightarrow \quad 1 = \lambda - \lambda\varphi'(x_0),$$

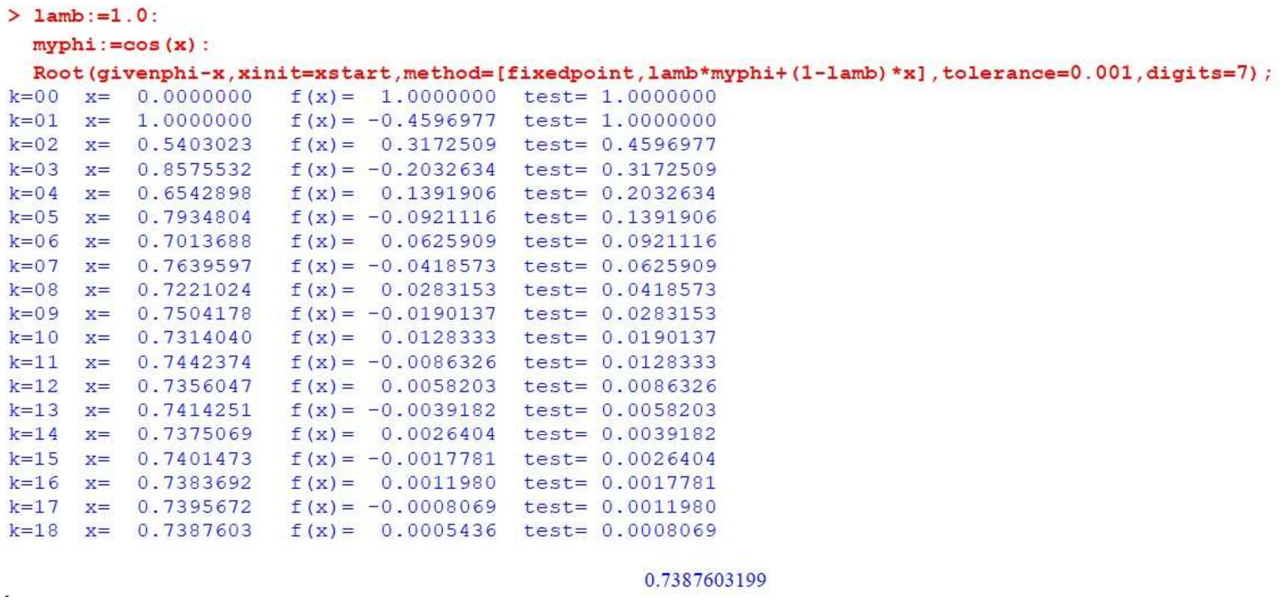
and therefore the **optimal** value of λ, denoted by λ_{opt}, near the point x_0 is given by

$$\lambda_{\text{opt}} = \frac{1}{1 - \varphi'(x_0)}. \tag{26}$$

Example 23.1. Consider once again the equation $\cos(x) = x$. In this case, we have $\varphi(x) = \cos(x)$. If we apply (26) at $x_0 = 0$, then

$$\lambda_{\text{opt}} = \frac{1}{1 + \sin(x_0)} = \frac{1}{1 + \sin(0)} = 1,$$

which does not seem to perform well (see Figure 115).



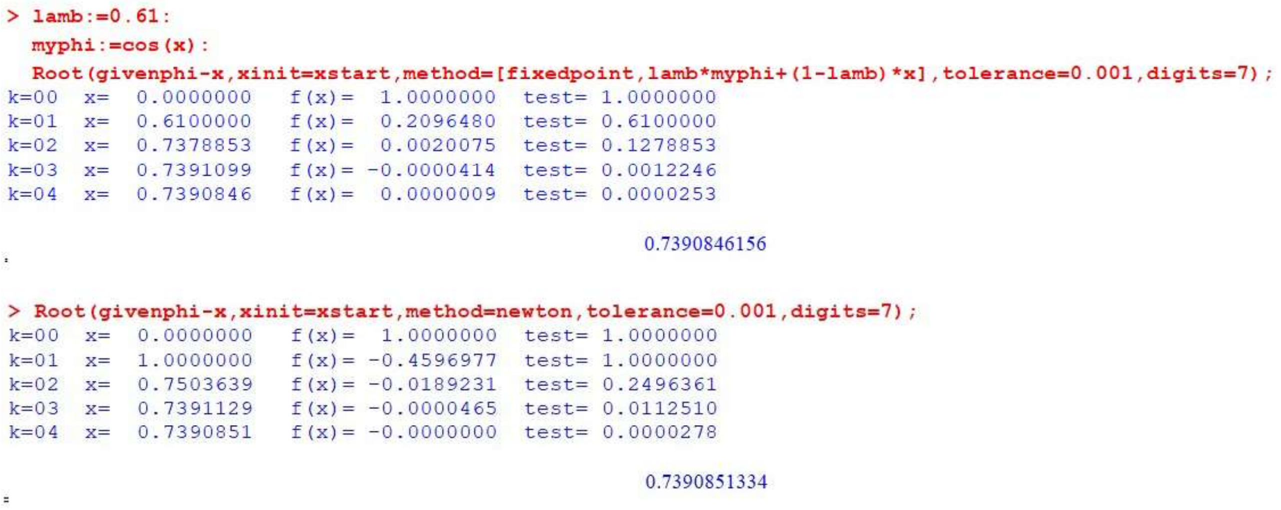
```
> lamb:=1.0:
  myphi:=cos(x):
  Root(givenphi-x,xinit=xstart,method=[fixedpoint,lamb*myphi+(1-lamb)*x],tolerance=0.001,digits=7);
k=00  x=  0.0000000   f(x)=  1.0000000  test= 1.0000000
k=01  x=  1.0000000   f(x)= -0.4596977  test= 1.0000000
k=02  x=  0.5403023   f(x)=  0.3172509  test= 0.4596977
k=03  x=  0.8575532   f(x)= -0.2032634  test= 0.3172509
k=04  x=  0.6542898   f(x)=  0.1391906  test= 0.2032634
k=05  x=  0.7934804   f(x)= -0.0921116  test= 0.1391906
k=06  x=  0.7013688   f(x)=  0.0625909  test= 0.0921116
k=07  x=  0.7639597   f(x)= -0.0418573  test= 0.0625909
k=08  x=  0.7221024   f(x)=  0.0283153  test= 0.0418573
k=09  x=  0.7504178   f(x)= -0.0190137  test= 0.0283153
k=10  x=  0.7314040   f(x)=  0.0128333  test= 0.0190137
k=11  x=  0.7442374   f(x)= -0.0086326  test= 0.0128333
k=12  x=  0.7356047   f(x)=  0.0058203  test= 0.0086326
k=13  x=  0.7414251   f(x)= -0.0039182  test= 0.0058203
k=14  x=  0.7375069   f(x)=  0.0026404  test= 0.0039182
k=15  x=  0.7401473   f(x)= -0.0017781  test= 0.0026404
k=16  x=  0.7383692   f(x)=  0.0011980  test= 0.0017781
k=17  x=  0.7395672   f(x)= -0.0008069  test= 0.0011980
k=18  x=  0.7387603   f(x)=  0.0005436  test= 0.0008069

                                      0.7387603199
```

Figure 115: The fixed-point method (relaxation) with $\lambda_{\text{opt}} = 1$.

On the other hand, if we start with $x_0 = 0.7$, then we obtain

$$\lambda_{\text{opt}} = \frac{1}{1+\sin(0.7)} \approx 0.61$$

and in this case Maple requires only 4 iterations (see Figure 116), which matches the number of iterations needed by Newton's method (also in Figure 116).



```
> lamb:=0.61:
  myphi:=cos(x):
  Root(givenphi-x,xinit=xstart,method=[fixedpoint,lamb*myphi+(1-lamb)*x],tolerance=0.001,digits=7);
k=00  x=  0.0000000   f(x)=  1.0000000  test= 1.0000000
k=01  x=  0.6100000   f(x)=  0.2096480  test= 0.6100000
k=02  x=  0.7378853   f(x)=  0.0020075  test= 0.1278853
k=03  x=  0.7391099   f(x)= -0.0000414  test= 0.0012246
k=04  x=  0.7390846   f(x)=  0.0000009  test= 0.0000253

                                      0.7390846156

> Root(givenphi-x,xinit=xstart,method=newton,tolerance=0.001,digits=7);
k=00  x=  0.0000000   f(x)=  1.0000000  test= 1.0000000
k=01  x=  1.0000000   f(x)= -0.4596977  test= 1.0000000
k=02  x=  0.7503639   f(x)= -0.0189231  test= 0.2496361
k=03  x=  0.7391129   f(x)= -0.0000465  test= 0.0112510
k=04  x=  0.7390851   f(x)= -0.0000000  test= 0.0000278

                                      0.7390851334
```

Figure 116: The fixed-point method (relaxation) with $\lambda_{\text{opt}} \approx 0.61$.

In the graph in Figure 117 we can see the comparison between the $\cos(x)$ function (in blue) and the graph of $\lambda_{0.61}(x)$ in the green. The figure shows that the relaxed fixed-point function is substantially flatter than the original map $\varphi(x)$ on the interval under consideration, especially near the intersection with the line $y = x$, which represents the fixed point. This is precisely the desired effect of relaxation: instead of working with the steeper curve $\varphi(x)$, we replace it with a function whose slope has smaller magnitude around the solution, thereby making the fixed-point iteration more stable and, in principle, faster to converge. In particular, the green curve appears to be almost horizontal near the fixed point, which suggests that the choice of λ is effective.

```
> plot([x,givenphi,lamb*givenphi+(1-lamb)*x],x=0..1,scaling=constrained,color=[red,navy,green],legend=["x",
  "phi","relaxed phi"]);
```

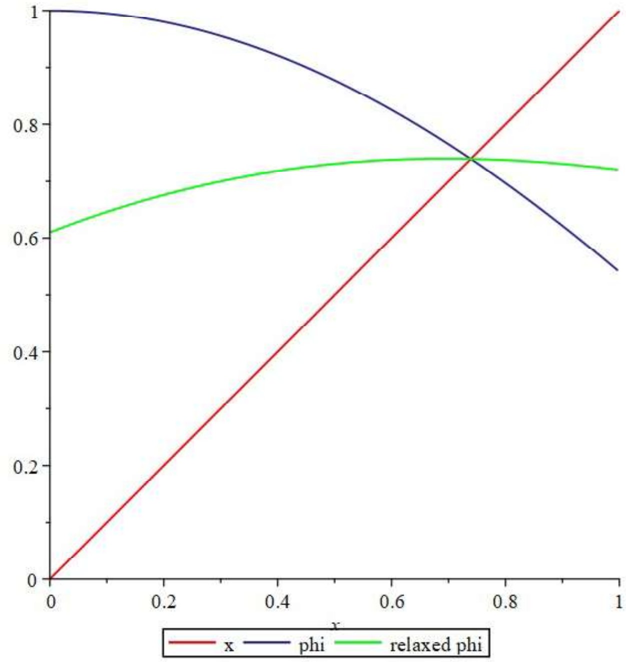


Figure 117: The optimal $\lambda_{\text{opt}} \approx 0.61$.

23.2 Root finding through fixed points (relaxation for root problems)

Suppose we are once again trying to solve an equation of the form $f(x) = 0$. As we have seen, we may use the bisection, Newton, or secant methods. Let us now consider a different approach based on the fixed-point method. Starting from the equation $f(x) = 0$, we add x to both sides and obtain

$$f(x) + x = x.$$

If we now define $\varphi(x) := f(x) + x$, then we obtain a fixed-point problem. We will apply relaxation. In this case, we have

$$\varphi_\lambda(x) = \lambda(f(x) + x) + (1 - \lambda)x = \lambda f(x) + x.$$

Assume now that we have already applied some iterations to this function φ_λ and obtained a point x_k. How can we find an optimal value of λ around x_k, denoted by $\lambda_{\text{opt},k}$? We apply formula (26) at x_k as before and obtain

$$\lambda_{\text{opt},k} = \frac{1}{1 - \varphi'(x_k)} = \frac{1}{1 - (f'(x_k) + 1)} = -\frac{1}{f'(x_k)}.$$

Now the iteration formula becomes

$$x_{k+1} = \varphi_{\lambda_{\mathrm{opt},k}}(x_k) = -\frac{1}{f'(x_k)}\, f(x_k) + x_k.$$

In other words,

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}$$

which is exactly Newton's method. This perspective helps explain the remarkable efficiency of Newton's method: rather than being viewed merely as a root-finding algorithm, it can be understood as a relaxed fixed-point iteration in which the parameter λ is chosen optimally at every step. In this sense, Newton's method is not just fast by accident; its speed comes from adapting the iteration so that the associated fixed-point map is as flat as possible near the current iterate.

23.3 Review of analytic methods: recognizing the appropriate technique

Before solving a differential equation, one of the most important steps is to recognize its form and decide which method is appropriate. In this course, there are three analytic methods that we have emphasized: separation of variables the method of guessing undetermined coefficients, and variation of parameters. Each of them applies only under specific structural conditions. A first-order equation is a good candidate for **separation of variables** when it can be written in the form $y' = f(x)g(y)$, so that all the terms involving y can be moved to one side and all the terms involving x to the other. The method of **guessing** is used for linear equations with constant coefficients and a special right-hand side, such as polynomials, exponentials, sines, cosines, or products of these. Finally, **variation of parameters** applies to linear equations, either first order or higher order, and is particularly useful when the equation is not separable and the right-hand side is not of the special form required for guessing. The following examples are intended as a review of how to recognize which method can be used in principle.

Example 23.2. Consider the differential equation

$$y' = \frac{x^2+1}{y^2+1}.$$

This is a first-order equation, and it is naturally written as

$$(y^2+1)\, y' = x^2 + 1.$$

Hence, it can be rewritten in separated form as

$$(y^2+1)\, dy = (x^2+1)\, dx.$$

Therefore, the correct method here is separation of variables.

Example 23.3. Consider the differential equation

$$y'' - 4y' + 13y = e^{2x}\cos(3x).$$

This is a linear second-order equation with constant coefficients. Moreover, the right-hand side is of a very special type, namely an exponential times a trigonometric function. For equations of this form, the appropriate method is guessing (also called the method of undetermined coefficients). Indeed, one looks for a particular solution of the form

$$y_p = e^{2x}\Big(A\cos(3x) + B\sin(3x)\Big).$$

Thus, this is a standard example where guessing can be used in principle.

Example 23.4. Consider the differential equation

$$y' + \frac{2}{x} y = \ln(x), \qquad x > 0.$$

This is a linear first-order equation, so it is not treated by guessing, since that method is designed for linear equations with constant coefficients and special right-hand sides. It is also not naturally separable, because the terms involving y and the terms involving x cannot be split into a product of the form $f(x)g(y)$. Therefore, the appropriate method here is variation of parameters in the first-order setting. The key point is that the equation is linear, but neither separable nor of the special constant-coefficient type required for guessing.

Example 23.5. Consider the differential equation

$$y'' + y = \tan(x).$$

This is a linear second-order equation with constant coefficients, but the right-hand side $\tan(x)$ is not one of the standard forcing terms for which guessing works well. In particular, it is not a polynomial, not a pure exponential, not a sine or cosine alone, and not a finite combination of the usual types used in undetermined coefficients. Therefore, the method of guessing is not appropriate here. The equation is also clearly not separable, since it is of second order and linear. Hence, the correct method to use in principle is variation of parameters. This is precisely the kind of example where variation of parameters becomes necessary because the equation is linear, but the forcing term is not suited to guessing.

23.4 Review of eigenvalues and eigenvectors

Let us briefly review how to compute eigenvalues and eigenvectors of a matrix as we will need it soon. Recall that if A is a square matrix, a scalar λ is called an **eigenvalue** of A if there exists a nonzero vector v such that $Av = \lambda v$. Any nonzero vector v satisfying this equation is called an **eigenvector** associated with λ. In practice, to find the eigenvalues of a matrix A, we solve the characteristic equation

$$\det(A - \lambda I) = 0.$$

Once the eigenvalues are known, we find the corresponding eigenvectors by solving the linear system $(A - \lambda I)v = 0$. We now review this procedure with some examples.

Example 23.6. Consider the matrix $A = \begin{pmatrix} 7 & -6 \\ 6 & -6 \end{pmatrix}$. To find the eigenvalues, we compute the determinant

$$\det(A - \lambda I) = \det \begin{pmatrix} 7 - \lambda & -6 \\ 6 & -6 - \lambda \end{pmatrix} = \lambda^2 - \lambda - 6.$$

Thus, the characteristic equation is $\lambda^2 - \lambda - 6 = 0$, so the eigenvalues are $\lambda = -2$ and $\lambda = 3$. For $\lambda = -2$, we solve $(A + 2I)v = 0$. This leads to the relation $3v_1 - 2v_2 = 0$, so we may choose $v_1 = 2$ and $v_2 = 3$. Therefore, one eigenvector associated with $\lambda = -2$ is $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$. For $\lambda = 3$, we solve $(A - 3I)v = 0$. This gives $2v_1 - 3v_2 = 0$, and hence we may take $v_1 = 3$ and $v_2 = 2$. Thus, one eigenvector associated with $\lambda = 3$ is $\begin{pmatrix} 3 \\ 2 \end{pmatrix}$.

Example 23.7. Consider now the matrix $A = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}$. Its characteristic polynomial is

$$\det(A - \lambda I) = \det \begin{pmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{pmatrix} = (2-\lambda)^2 - 1.$$

Hence, the characteristic equation becomes $\lambda^2 - 4\lambda + 3 = 0$, so the eigenvalues are $\lambda = 1$ and $\lambda = 3$. For $\lambda = 1$, we solve $(A - I)v = 0$. This gives $v_1 + v_2 = 0$, so one possible eigenvector is $\begin{pmatrix} 1 \\ -1 \end{pmatrix}$. For $\lambda = 3$, we solve $(A - 3I)v = 0$. This gives $-v_1 + v_2 = 0$, so one possible eigenvector is $\begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Example 23.8. Let us now consider the matrix $A = \begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix}$. We compute

$$\det(A - \lambda I) = \det \begin{pmatrix} 1-\lambda & 2 \\ 4 & 5-\lambda \end{pmatrix} = \lambda^2 - 6\lambda + 13.$$

Therefore, the characteristic equation is $\lambda^2 - 6\lambda + 13 = 0$, and solving it we obtain the eigenvalues $\lambda = 3 \pm 2i$. Let us find an eigenvector corresponding to $\lambda = 3 + 2i$. We solve $(A - \lambda I)v = 0$, that is,

$$\begin{pmatrix} 1-(3+2i) & 2 \\ 4 & 5-(3+2i) \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

This becomes

$$\begin{pmatrix} -2-2i & 2 \\ 4 & 2-2i \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}.$$

From the first row, we get $(-2-2i)v_1 + 2v_2 = 0$, or equivalently $-(1+i)v_1 + v_2 = 0$. If we choose $v_1 = 1 - i$, then $v_2 = (1+i)(1-i) = 2$. Thus, one eigenvector associated with $\lambda = 3 + 2i$ is $\begin{pmatrix} 1-i \\ 2 \end{pmatrix}$. Similarly, for $\lambda = 3 - 2i$, one finds an eigenvector $\begin{pmatrix} 1+i \\ 2 \end{pmatrix}$.