Green Functions: The Bridge Between Linear Algebra and Differential Equations

Andrew Carnes

October 26, 2018

Contents

1   INTRODUCTION

This document is a real adventure. We look at functions as vectors and reinterpret derivatives as matrices. By viewing derivatives as matrices, we can view differential equations as linear algebra problems and use the techniques from linear algebra to solve differential equations (Diff EQ). The note starts by extrapolating dot products from vectors to functions, then uses dot products to develop matrix multiplication on functions, and then uses the matrix on function action to make the connection between linear algebra and Diff EQ. Looking at Diff EQ in this light, we find that the Green function is the inverse of a differential operator.

To provide some initial insight, consider the linear algebra problem $Mv = b$ with the matrix M and the vectors v and b. Finding M^{-1} solves for v via $v = M^{-1}b$. In the same way, the Green function $G = D^{-1}$ solves the differential equation $Df = g$ via $\int G(x,y)g(y)dy = f(x)$. The journey from vectors to functions and linear algebra to Diff EQ provides some intersting insight into some unsuspecting places like the Dirac delta function, the fundamental theorem of calculus, and Fourier transforms.

Looking at Green functions this way finally made them click and helped me a bunch with electromagnetism and quantum field theory in graduate school. Nothing here is super rigorous, but it gets the point across and provides a new way to look at some familiar things. Anywho, let's get on with the adventure.

2   FUNCTIONS AS VECTORS

This section compares functions to vectors and extrapolates the familiar vector dot product to functions, providing the foundation of the framework that takes us from matrices and vectors to functions and Diff EQ. Onwards! What is a vector? A vector $\vec{v}$ is an ordered 1D storage device. A vector of length N has N bins to put numbers in, and you keep track of the bins by labeling them with an integer. Note that you can only put one number in each bin. If we label the bins with i, v_i is the information in the ith bin. $\vec{v}$ usually written like so,

$$\vec{v} = \begin{pmatrix} v_1 \\ v_2 \\ ... \\ v_N \end{pmatrix}. \tag{1}$$

Now consider a function. A function f(x) is an ordered 1D storage device with a bunch of bins. The bins are labeled by x, and f(x) is the number in the xth bin. Each bin only has one number. In this case, x is a real number so f(x) is much more dense than $\vec{v}$. f(x) can be

written like so,

$$f(x) = \begin{pmatrix} f(-\infty) \\ ... \\ f(-3.11111111128) \\ ... \\ f(-3.1111111112) \\ ... \\ f(0) \\ ... \\ f(912.99) \\ ... \\ f(999999999) \\ ... \\ f(\infty) \end{pmatrix}. \quad (2)$$

Between any two bins there are an infinite number of bins (welcome to the real numbers), which makes things tricky, but aside from the storage density, functions and vectors work the same way. Give them a location and they spit out the info you want.

Let's discretize the real number line into chunks of width dx. We can then label the entries in f(x) by $f(x_i)$, where $x_{i+1} - x_i = dx$. Then we have

$$f(x) = \begin{pmatrix} f(x_{-\infty}) \\ ... \\ ... \\ f(x_{-2}) \\ ... \\ f(x_{-1}) \\ ... \\ f(x_0) \\ ... \\ f(x_1) \\ ... \\ f(x_2) \\ ... \\ ... \\ f(x\infty) \end{pmatrix}. \quad (3)$$

Note that there are still an infinite number of entries between $f(x_i)$ and $f(x_{i+1})$. Now let's take dot products. With two vectors, $\vec{v}$ and $\vec{u}$, the dot product is

$$\vec{v} \cdot \vec{u} = v_1 u_1 + v_2 u_2 + ... v_N u_N = \sum_{i=1}^{N} v_i u_i. \quad (4)$$

As usual, when $\vec{v} \cdot \vec{u} = 0$ the vectors are orthogonal, and when $\frac{\vec{v} \cdot \vec{u}}{\sqrt{v^2 u^2}} = 0$ the vectors are parallel. Moreover, when all v_i and u_i are 1, $\vec{v} \cdot \vec{u} = N$, the total number of components.

4

If we consider only a subset of components with all $v_i = u_i = 1$ then $\sum_{i=1}^{N_c} v_i u_i = N_c$, the number of components summed over.

Now take a dot product of functions f(x) and g(x),

$$
\begin{aligned}
f(x) \cdot g(x) =& ... + f(x_{-2})g(x_{-2}) + ... + f(x_{-1})g(x_{-1}) \\
& + ... + f(x_0)g(x_0) + ... + f(x_1)g(x_1) + ... + f(x_2)g(x_2) + ...
\end{aligned}
\tag{5}
$$

Let's look at the sum from x_1 to x_2 when f(x) = g(x) = 1 in all bins

$$
f(x) \cdot g(x) = f(x_1)g(x_1) + ... + ... \text{ (up to } f(x_2)g(x_2)) = 1 + 1 + ... + 1 = N_c.
\tag{6}
$$

This gives us the number of components summed over, N_c, which is infinite. That's true, but it's not a helpful normalization. Let's redefine the dot product between functions to yield the size of the real interval summed over, $x_2 - x_1 = dx$.

$$
\begin{aligned}
f(x) \cdot g(x) &= f(x_1)g(x_1)\epsilon + ... + ... \text{ (up to } f(x_2)g(x_2)) \\
&= 1\epsilon + ... + 1\epsilon = N_c\epsilon = dx.
\end{aligned}
\tag{7}
$$

The full dot product is now

$$
\begin{aligned}
f(x) \cdot g(x) =& ... + f(x_{-2})g(x_{-2})\epsilon + ... + f(x_{-1})g(x_{-1})\epsilon \\
& + ... + f(x_0)g(x_0)\epsilon + ... + f(x_1)g(x_1)\epsilon + ... + f(x_2)g(x_2)\epsilon + ...
\end{aligned}
\tag{8}
$$

Finally, let's represent all N_c terms between f(x_i)g(x_i) and f(x_{i+1})g(x_{i+1}) by f(x_i)g(x_i) using the assumption that f(x) and g(x) change very little throughout the interval dx. This yields

$$
\begin{aligned}
f(x) \cdot g(x) =& ... + f(x_{-2})g(x_{-2})N_c\epsilon + f(x_{-1})g(x_{-1})N_c\epsilon + f(x_0)g(x_0)N_c\epsilon \\
& + f(x_1)g(x_1)N_c\epsilon + f(x_2)g(x_2)N_c\epsilon + ...
\end{aligned}
\tag{9}
$$

Remembering that $N_c\epsilon = dx$, we get

$$
f(x) \cdot g(x) = \sum_{i=-\infty}^{\infty} f(x_i)g(x_i)dx = \int_{-\infty}^{\infty} f(x)g(x)dx.
\tag{10}
$$

As with vectors, when $f(x) \cdot g(x) = 0$, two functions are orthogonal.

Now that we have dot products between functions down, we can develop matrix multiplication on functions. We just view matrix multiplication in terms of dot products, and BAM we've got ourselves some wacky version of linear algebra. Let's move on to this next part.

3 FROM MATRICES TO 2D FUNCTIONS

This section develops matrix multiplication on functions in terms of dot products. Once we have this down, we can move on to write the derivative as a matrix and develop the connection between linear algebra and Diff EQ. OK so, let's figure out this matrix-function business.

A matrix transforms a vector into another vector. Consider a rotation,

$$R\vec{v} = \begin{pmatrix} cos\theta & sin\theta \\ -sin\theta & cos\theta \end{pmatrix} \begin{pmatrix} v_1 \\ v_2 \end{pmatrix} = \begin{pmatrix} v_1 cos\theta + v_2 sin\theta \\ -v_1 sin\theta + v_2 cos\theta \end{pmatrix} = \vec{v'}. \tag{11}$$

The vector $\vec{v'}$ is still a vector, it's just $\vec{v}$ rotated by θ. We can write the matrix multiplication in summation form like so $v_i' = \sum_j R_{ij} v_j$. We can also write the transformation in terms of dot products $v_i' = \vec{R}_i \cdot \vec{v}$, where $\vec{R}_i$ is the row vector in row i of the rotation matrix R.

We can use the function dot product to extend this to functions, $g(x) = K(x) \cdot f$, where K(x) is the $x^t h$ row of the matrix K. Writing out the dot product in terms of an integral yields $\int K(x,y)f(y)dy$. Integrating f(y) against the 2D function K(x,y) performs the matrix multiplication yielding the new function g(x). Let's write $Kf = g$ to represent this operation. Imagine the discretized matrix multiplication,

$$\begin{pmatrix} K(y_1, x_1) & K(y_1, x_2) & ... & K(y_1, x_N) \\ K(y_2, x_1) & K(y_2, x_2) & ... & K(y_2, x_N) \\ ... & ... & ... & ... \\ K(y_N, x_1) & K(y_N, x_2) & ... & K(y_N, x_N) \end{pmatrix} \begin{pmatrix} f(x_1) \\ f(x_2) \\ ... \\ f(x_N) \end{pmatrix}$$

$$= \begin{pmatrix} K(y_1, x_1)f(x_1)dx + K(y_1, x_2)f(x_2)dx + ... + K(y_1, x_N)f(x_N) \\ K(y_2, x_1)f(x_1)dx + K(y_2, x_2)f(x_2)dx + ... + K(y_2, x_N)f(x_N) \\ ... + ... + ... + ... \\ K(y_N, x_1)f(x_1)dx + K(y_N, x_2)f(x_2)dx + ... + K(y_N, x_N)f(x_N) \end{pmatrix} = \begin{pmatrix} g(y_1) \\ g(y_2) \\ ... \\ g(y_N) \end{pmatrix}. \tag{12}$$

An example is the Fourier transform,

$$F(k,x)f(x) = \frac{1}{2\pi} \int e^{-ikx} f(x)dx = g(k), \tag{13}$$

which we will come back to later.

4   THE DERIVATIVE AS A MATRIX

Just like a matrix turns a vector into another vector, the derivative turns a function into another function, $\frac{d}{dx}f(x) = \frac{f(x+dx)-f(x)}{dx} = f'(x)$. If we can understand the derivative as a matrix, we can turn differential equations (Diff EQ) problems into linear algebra problems and solve them using the same techniques.

Let's write $\frac{d}{dx}f(x)$ in our matrix multiplication $Df = \int D(y,x)f(x)dx = g(y)$ formalism,

$$Df = \frac{1}{dxdx}\begin{pmatrix} \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ \cdots & -1 & 1 & 0 & 0 & 0 & 0 & \cdots \\ \cdots & 0 & -1 & 1 & 0 & 0 & 0 & \cdots \\ \cdots & 0 & 0 & -1 & 1 & 0 & 0 & \cdots \\ \cdots & 0 & 0 & 0 & -1 & 1 & 0 & \cdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \end{pmatrix}\begin{pmatrix} \cdots \\ f(x_i) \\ f(x_{i+1}) \\ f(x_{i+2}) \\ f(x_{i+3}) \\ f(x_{i+4}) \\ f(x_{i+5}) \\ \cdots \end{pmatrix} = \begin{pmatrix} \cdots \\ \frac{-f(x_i)+f(x_{i+1})}{dxdx}dx \\ \frac{-f(x_{i+1})+f(x_{i+2})}{dxdx}dx \\ \frac{-f(x_{i+2})+f(x_{i+3})}{dxdx}dx \\ \frac{-f(x_{i+3})+f(x_{i+4})}{dxdx}dx \\ \frac{-f(x_{i+4})+f(x_{i+5})}{dxdx}dx \\ \frac{-f(x_{i+5})+f(x_{i+6})}{dxdx}dx \\ \cdots \end{pmatrix}$$

$$= \begin{pmatrix} \cdots \\ \frac{f(x_{i+1})-f(x_i)}{dx} \\ \frac{f(x_{i+2})-f(x_{i+1})}{dx} \\ \frac{f(x_{i+3})-f(x_{i+2})}{dx} \\ \frac{f(x_{i+4})-f(x_{i+3})}{dx} \\ \frac{f(x_{i+5})-f(x_{i+4})}{dx} \\ \frac{f(x_{i+6})-f(x_{i+5})}{dx} \\ \cdots \end{pmatrix} = \begin{pmatrix} \cdots \\ g(y_i) \\ g(y_{i+1}) \\ g(y_{i+2}) \\ g(y_{i+3}) \\ g(y_{i+4}) \\ g(y_{i+6}) \\ \cdots \end{pmatrix}$$

$$(14)$$

Note that an integral multiplies by a factor of dx, and a derivative divides by dx. We need $\frac{d}{dx}f = \int D(y,x)f(x)dx$ with dimensions like $\frac{1}{dx} \sim Ddx$; therefore, $D \sim \frac{1}{dxdx}$.

So it turns out you can pretty much just write down the derivative in matrix form. How nice. Let's use this to check out the fundamental theorem of calculus, and then return to the whole differential equations are just linear algebra business.

5   THE FUNDAMENTAL THEOREM OF CALCULUS

The matrixified derivative provides some interesting insight into a familiar theorem of calculus. If I remember right, the fundamental theorem of calculus is $\int_a^b \frac{d}{dx}f(x)dx = f(a) - f(b)$.

Let's see what happens if we integrate $\frac{d}{dx}f(x)$.

$$\int_a^b f'(x)dx = 1 \cdot f'(x) = dx \begin{pmatrix} \dots & 1 & 1 & 1 & 1 & 1 & 1 & \dots \end{pmatrix} \begin{pmatrix} \dots \\ \frac{f(x_{i+1})-f(x_i)}{dx} \\ \frac{f(x_{i+2})-f(x_{i+1})}{dx} \\ \frac{f(x_{i+3})-f(x_{i+2})}{dx} \\ \frac{f(x_{i+4})-f(x_{i+3})}{dx} \\ \frac{f(x_{i+5})-f(x_{i+4})}{dx} \\ \frac{f(x_{i+6})-f(x_{i+5})}{dx} \\ \dots \end{pmatrix}$$

$$= f(b) + \dots + f(x_{i+6}) - f(x_{i+5}) + f(x_{i+5}) - f(x_{i+4}) + f(x_{i+4}) - f(x_{i+3}) + \dots - f(a)$$
$$= f(b) - f(a)$$

$$(15)$$

Boom! The fundamental theorem of calculus. Nailed it.

6   LINEAR ALGEBRA IN A NUTSHELL

Now onto linear algebra. If the derivative is a matrix, then any differential equation is just some linear algebra problem. Let's figure out a general method to solve some linear algebra mystery and then we can use the same strategy for Diff EQ. With a linear algebra problem like $M\vec{v} = \vec{b}$, we can solve for $\vec{v}$ by finding M^{-1}. With M^{-1} in hand, $\vec{v} = M^{-1}\vec{b}$. Problem solved. But how do we find M^{-1}? Well, M^{-1} is easy to find when M is diagonal. Let's go to the bar basis where M is diagonal.

$$\bar{M}^{-1}\bar{M} = 1 = \begin{pmatrix} \dots & \dots & \dots & \dots & \dots \\ \dots & 1/m_i & 0 & 0 & \dots \\ \dots & 0 & 1/m_{i+1} & 0 & \dots \\ \dots & 0 & 0 & 1/m_{i+2} & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix} \begin{pmatrix} \dots & \dots & \dots & \dots & \dots \\ \dots & m_i & 0 & 0 & \dots \\ \dots & 0 & m_{i+1} & 0 & \dots \\ \dots & 0 & 0 & m_{i+2} & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

$$\to \bar{M}^{-1} = \begin{pmatrix} \dots & \dots & \dots & \dots & \dots \\ \dots & 1/m_i & 0 & 0 & \dots \\ \dots & 0 & 1/m_{i+1} & 0 & \dots \\ \dots & 0 & 0 & 1/m_{i+2} & \dots \\ \dots & \dots & \dots & \dots & \dots \end{pmatrix}$$

$$(16)$$

In the bar basis, $\bar{M}$ consists of its eigenvalues along the diagonal, and $\bar{M}^{-1}$ is just one over the eigenvalues along the diagonal. The equation to solve becomes $\vec{\bar{v}} = \bar{M}^{-1}\vec{\bar{b}}$. If we assume there is some transformation matrix T that transforms the non-bar basis to the bar basis, then $T\vec{v} = \bar{M}^{-1}T\vec{b}$ and $\vec{v} = T^{-1}\bar{M}^{-1}T\vec{b}$. Finally,

$$M^{-1} = T^{-1}\bar{M}^{-1}T. \tag{17}$$

Almost there, we still have to find T and T^{-1}! $\vec{\bar{m}}_i$, the eigenvector in the bar basis corresponding to eigenvalue m_i, is just $\begin{pmatrix} \dots & 0 & 1 & 0 \dots \end{pmatrix}$, where the 1 is in the ith position. T^{-1}

8

converts from the bar basis to the non-bar basis and should convert $\vec{\bar{m}}_i$ to $\vec{m}_i$. This holds for all eigenvectors. Naturally,

$$T^{-1}\vec{\bar{m}}_i = \vec{m}_i \to T^{-1} = \left(\ \vec{m}_1 \mid \vec{m}_2 \mid \ ... \mid \vec{m}_i \mid \ ... \mid \vec{m}_N \ \right) \tag{18}$$

Note that $\vec{m}_i$ defines the entire i^{th} column of the T^{-1} matrix. Checking this,

$$T^{-1}\vec{\bar{m}}_i = \left(\ \vec{m}_1 \mid \vec{m}_2 \mid \ ... \mid \vec{m}_i \mid \ ... \mid \vec{m}_N \ \right) \begin{pmatrix} 0 \\ 0 \\ ... \\ 1 \\ ... \\ 0 \end{pmatrix} = 0\vec{v}_1 + 0\vec{v}_2 + ... + \vec{v}_i + ... + 0\vec{v}_N = \vec{v}_i. \tag{19}$$

Our transformation consists of orthonormal eigenvectors which implies that $T^\dagger T = 1$. Note that $\dagger$ represents the Hermitian conjugate.

$$T^\dagger T = \begin{pmatrix} \vec{v}_1^\dagger \vec{v}_1 & \vec{v}_1^\dagger \vec{v}_2 & ... & \vec{v}_1^\dagger \vec{v}_N \\ \vec{v}_2^\dagger \vec{v}_1 & \vec{v}_2^\dagger \vec{v}_2 & ... & \vec{v}_2^\dagger \vec{v}_N \\ ... & ... & ... & ... \\ \vec{v}_N^\dagger \vec{v}_1 & \vec{v}_N^\dagger \vec{v}_2 & ... & \vec{v}_N^\dagger \vec{v}_N \end{pmatrix} = 1 \tag{20}$$

$T^\dagger T = 1$ further implies that $T^{-1} = T^\dagger$ and $T = T^{-1\dagger}$. We now have the entire solution in hand. $M^{-1} = T^{-1}\bar{M}^{-1}T = T^\dagger \bar{M}^{-1}T$ with $\bar{M}$ given by one over the eigenvalues of M (Equation 16) and $T^{-1} = T^\dagger$ given by the eigenvectors of M (Equation 18).

Let's go through a simple example to solidify things.