Functional Calculus for Matrices.

Alonso Delfín
University of Oregon.

February 22, 2018

Abstract.

Below I will explain how to make sense of $f(T)$ where T is an $n \times n$ complex matrix and f is a $C^M(U)$ function, where the integer M and the set U will be determined by the matrix T. I will be using elementary linear algebra techniques, so this should be accessible to everyone. I will cover an example and some general properties of this functional calculus.

Time permitting, I'll explain how this is related to functional analysis: turns out this functional calculus gives rise to a projection valued discrete measure, so this is in fact a particular case of the Borel functional calculus.

Goal.

Consider the Banach Algebra $\mathbb{C}^n$. Then, $\mathcal{B}(\mathbb{C}^n) = M_n(\mathbb{C}^n)$. Fix $T \in M_n(\mathbb{C}^n)$ and take p a polynomial in $\mathbb{C}$. We may define $p(T) \in M_n(\mathbb{C}^n)$ in the obvious way. Our goal is to extend the map $p \mapsto p(T)$ to a bigger class of functions containing the polynomials in $\mathbb{C}$.

Background and main definition.

We fix $T \in M_n(\mathbb{C}^n)$ and consider $\sigma(T) := \{\lambda_1, \dots, \lambda_k\}$ to be the set of distinct eigenvalues of T. If each λ_j has multiplicity $\mathrm{mult}(\lambda_j)$, then we have

$$n = \sum_{j=1}^{k} \mathrm{mult}(\lambda_j)$$

Now let ψ_T be the minimal polynomial of T, that is the monic polynomial of least degree such that when evaluating T we get the zero matrix. It's well known that

$$\psi_T(x) = \prod_{j=1}^{k} (x - \lambda_j)^{m_j} \tag{1}$$

where $1 \leq m_j \leq \text{mult}(\lambda_j)$. Let m be the degree of ψ_T. Let χ_T be the characteristic polynimial of T, that is $\chi_T(x) := \det(T - xI)$. Then, by Cayley-Hamilton theorem, one has $\chi_T(T) = 0$. So, we have that χ_T is a multiple of ψ_T.

Lemma. *If g and h are non-zero polynomials, then $g(T) = h(T)$ if and only if*

$$g(\lambda_j) = h(\lambda_j),\ g'(\lambda_j) = h'(\lambda_j),\ \ldots,\ g^{(m_j-1)}(\lambda_j) = h^{(m_j-1)}(\lambda_j) \tag{2}$$

for all $j = 1, \ldots, k$.

Proof. Suppose first that g and h are non-zero polynomials for which we have $g(T) = h(T)$. Then, if $d := g - h$ we have $d(T) = 0$. Hence, d must have degree greater or equal than m and must be a multiple of ψ_T. That is, there is a polynomial q such that

$$d = \psi_T \cdot q$$

Hence, $d' = \psi_T \cdot q + \psi_T \cdot q'$. But, it follows from (1) that

$$d(\lambda_j) = 0,\ d'(\lambda_j) = 0,\ \ldots,\ d^{(m_j-1)}(\lambda_j) = 0$$

for all $j = 1, \ldots, k$, which in turn gives that

$$g(\lambda_j) = h(\lambda_j),\ g'(\lambda_j) = h'(\lambda_j),\ \ldots,\ g^{(m_j-1)}(\lambda_j) = h^{(m_j-1)}(\lambda_j)$$

for all $j = 1, \ldots, k$.

Conversely, if (2) holds, then, for each $j = 1, \ldots, k$, the polynomial $g - h$ has a zero of multiplicity a least m_j at λ_j. Thus, $g - h$ is a multiple of Ψ_T and therefore $(g - h)(T) = 0$. ■

The previous lemma implies that for any polynomial p, the matrix $p(T)$ depends only on the values

$$p(\lambda_j),\ p'(\lambda_j),\ \ldots,\ p^{(m_j-1)}(\lambda_j)$$

for $j = 1, \ldots, k$.

From now on we consider functions in $C^M(U)$ where $M := \max_j\{m_j - 1\}$ and U is an open subset of $\mathbb{C}$ such that $\sigma(T) \subset U$.

Definition. (Lagrange-Sylvester interpolation polynomial) For $f \in C^M(U)$, we define the polynomial ℓ_f by

$$\ell_f(x) := \sum_{j=1}^{k} \left(\sum_{i=0}^{m_j-1} f^{(i)}(\lambda_j) e_{j,i}(x) \right),$$

where each $e_{j,i}$ is a polynomial such that

$$e_{j,i}^{(l)}(\lambda_s) = \begin{cases} 1 & \text{if } j = s \text{ and } l = i \\ 0 & \text{otherwise} \end{cases}$$

▲

By construction, we get

$$\ell_f(\lambda_j) = f(\lambda_j),\ \ell_f'(\lambda_j) = f'(\lambda_j),\ \ldots,\ \ell_f^{(m_j-1)}(\lambda_j) = f^{(m_j-1)}(\lambda_j)$$

for all $j = 1, \ldots, k$. Thus, we define a functional calculus $f \mapsto f(T)$ by letting $f(T)$ be the matrix $\ell_f(T)$. Since ℓ_f is itself a polynomial and $\ell_p = p$ for any polynomial p, it follows that this functional calculus extends the polynomial one to the bigger class $C^M(U)$.

Example

Suppose that $T \in M_n(\mathbb{C})$ has ones above the diagonal ans zeros elsewhere. That is, T is given by

$$T = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ & \vdots & & \ddots & \vdots \\ 0 & 0 & 0 & \ldots & 1 \\ 0 & 0 & 0 & \ldots & 0 \end{pmatrix}$$

Notice that $\sigma(T) = \{0\}$ and that $\psi_T(x) = x^n$. Thus, for any $f \in C^{n-1}(U)$ (here U can be any open set containing 0), we have

$$\ell_f(x) = f(0) + f'(0)x + \frac{f''(0)}{2!}x^2 + \cdots + \frac{f^{(n-1)}(0)}{(n-1)!}x^{n-1}$$

Hence,

$$f(T) = \ell_f(T) = f(0)I + f'(0)T + \frac{f''(0)}{2!}T^2 + \cdots + \frac{f^{(n-1)}(0)}{(n-1)!}T^{n-1}.$$

Finally, we observe that multiplying T with itself j times, "pushes" the line of ones in the north east direction j times. This gives

$$f(T) = \begin{pmatrix} f(0) & f'(0) & \frac{f''(0)}{2!} & \cdots & \frac{f^{(n-1)}(0)}{(n-1)!} \\ 0 & f(0) & f'(0) & \cdots & \frac{f^{(n-2)}(0)}{(n-2)!} \\ \vdots & \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & \cdots & f(0) & f'(0) \\ 0 & 0 & \cdots & 0 & f(0) \end{pmatrix}$$

Properties of $f \mapsto f(T)$

Proposition. *Fix $T \in M_n(\mathbb{C})$. Let $f \in C^M(U)$, where M and U are as defined above.*

1. *If $S = P^{-1}TP$ for $P \in \mathrm{GL}_n(\mathbb{C})$, then $f(S) = P^{-1}f(T)P$.*
2. *If $T = \mathrm{diag}[T_1, \ldots, T_j]$, then $f(T) = \mathrm{diag}[f(T_1), \ldots, f(T_j)]$*

Proof. For *1.*, we know from basic linear algebra that S and T have the same minimal polynomial and same eigenvalues. Further, note that $S^k = P^{-1}T^kP$ for any integer k. Hence,

$$f(S) = \ell_f(S) = \ell_f(P^{-1}TP) = P^{-1}\ell_f(T)P = P^{-1}f(T)P$$

For *2.*, we notice first of all that, by the simplicity of the polynomial functional calculus, we have

$$f(T) = \ell_f(T) = \ell_f(\mathrm{diag}[T_1, \ldots, T_j]) = \mathrm{diag}[\ell_f(T_1), \ldots, \ell_f(T_j)]$$

So it suffices to prove that $\ell_f(T_k) = f(T_k)$. Indeed, observe that

$$\psi_T(\mathrm{diag}[0, \ldots, T_k, \ldots, 0]) = 0.$$

Thus f and $\ell_f^{(l)}$ agree on $\sigma(T_k)$. ■

One consequence of the previous proposition is the way the functional calculus behaves with diagonal matrices:

$$f\begin{pmatrix}\lambda_1 & & 0\\ & \ddots & \\ 0 & & \lambda_n\end{pmatrix} = \begin{pmatrix}f(\lambda_1) & & 0\\ & \ddots & \\ 0 & & f(\lambda_n)\end{pmatrix}$$

Theorem. *(Sylvester's Formula) Let T be a diagonalizable matrix, that is $T = P\Lambda P^{-1}$ for a diagonal matrix Λ. If U is an open set in $\mathbb{C}$ such that $\sigma(T) \subset U$ and $f \in C(U)$, then*

$$f(T) = \sum_{j=1}^{k} f(\lambda_j)P_j$$

where

$$P_j = \prod_{i\neq j} \frac{T-\lambda_i}{\lambda_j-\lambda_i}$$

Proof. Since T is a diagonalizable matrix, linear algebra tells us that if $\sigma(T) = \{\lambda_1, \dots, \lambda_k\}$, then

$$\psi_T(x) = \prod_{j=1}^{k}(x-\lambda_j)$$

Hence, ℓ_f is just the interpolation polynomial of f at $\sigma(T)$, otherwise simply known as the Lagrange polynomial. That is,

$$\ell_f(x) = \sum_{j=1}^{k} f(\lambda_j)\prod_{i\neq j}\left(\frac{x-\lambda_i}{\lambda_j-\lambda_i}\right)$$

Sylvester's formula now follows. ■

As a final remark, we show that Sylvester's Formula is a particular case of the Borel functional calculus. Indeed, notice that $P_iP_j = \delta_{i,j}P_j$ and that P_j is the projection onto the λ_j-eigenspace. So, if $\mathcal{P} : \mathfrak{B}(\mathbb{C}) \to M_n(\mathbb{C})$ is σ-additive, $\operatorname{supp}(\mathcal{P}) = \sigma(T)$ and $\mathcal{P}(\{\lambda_j\}) = P_j$. For each $z, w \in \mathbb{C}$, we have that $\mu_{z,w} : \mathfrak{B}(\mathbb{C}) \to \mathbb{C}$, given by $\mu_{z,w}(E) := \langle \mathcal{P}(E)z, w\rangle$ is a complex measure and that $\langle f(T)z, w\rangle = \int_{\sigma(T)} f d\mu_{z,w}$. So, in fact $f(T) = \int_{\sigma(T)} f d\mathcal{P}$.