Generalization of Sylvester formula

English mathematician James Joseph Sylvester derived his famous formula only for diagonal matrices. Another British mathematician, Arthur Buchheim, extended it for a general case of matrices reducible to Jordan form, e.g. those with some eigen values having multiplicity $l_i > 1$:

$$\det[\mathbf{M} - \lambda \mathbf{I}] = \prod_{i=1}^{m} \left(\lambda - \lambda_i\right)^{l_i}$$

I did not find derivation of Arthur Buchheim https://en.wikipedia.org/wiki/Sylvester%27s_formula and below is my own derivation that I found in my notes when I attempted to have complete set of matrices for accelerator quite long time ago…

Matrix functions and Projection operators

An arbitrary matrix **M** can be reduced to an unique matrix, which in general case has a Jordan form: for a matrix with arbitrary height of eigen values the set of eigen values $\{\lambda_1,...,\lambda_m\}$ contains only unique eigen values, i.e. $\lambda_k \neq \lambda_j;\ \forall\ \ k \neq j$:

$$size[\mathbf{M}] = M;\ \{\lambda_1,.....,\lambda_m\};\ m \le M;\ \det[\lambda_k \mathbf{I} - \mathbf{M}] = 0;$$

$$\mathbf{M} = \mathbf{U}\mathbf{G}\mathbf{U}^{-1};\ \ \mathbf{G} = \sum_{\oplus \mathbf{k=1,m}} \mathbf{G_k} = \mathbf{G_1} \oplus .... \oplus \mathbf{G_m};\ \ \sum size[\mathbf{G_k}] = M \tag{E-1}$$

where $\oplus$ means direct sum of block-diagonal square matrixes $\mathbf{G_k}$ which correspond to the eigen vector sub-space adjacent to the eigen value λ_k. Size of $\mathbf{G_k}$, which we call l_k, is equal to the multiplicity of the root λ_k of the characteristic equation

$$\det[\lambda \mathbf{I} - \mathbf{M}] = \prod_{k=1,m} (\lambda - \lambda_k)^{l_k}.$$

In general case, $\mathbf{G_k}$ is also a block diagonal matrix comprised of orthogonal sub-spaces belonging to the same eigen value

$$\mathbf{G_k} = \sum_{\oplus\, j=1,p_k} \mathbf{G_k^{\mathit{j}}} = \mathbf{G_1^1} \oplus .... \oplus \mathbf{G_m^{\mathit{p_k}}};\ \ \sum size[\mathbf{G_k^{\mathit{j}}}] = l_k \tag{E-2}$$

where we assume that we sorted the matrixes by increasing size: $size[\mathbf{G_k^{\mathit{j+1}}}] \ge size[\mathbf{G_k^{\mathit{j}}}]$, i.e. the

$$n_\mathbf{k} = size[\mathbf{G_k^{\mathit{p_k}}}] \le l_k \tag{E-3}$$

is the maximum size of the Jordan matrix belonging to the eigen value λ_k. General form of the Jordan matrix is:

$$\mathbf{G_k^n} = \begin{vmatrix} \lambda_{\mathbf{k}} & 1 & 0 & 0 \\ 0 & \lambda_{\mathbf{k}} & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & \lambda_{\mathbf{k}} \end{vmatrix} \qquad \text{(E-4)}$$

This is obviously including non-degenerate case when matrix **M** has M independent eigen values and all is just perfectly simple: matrix is reducible to a diagonal one

$$size[\mathbf{M}] = M; \ \{\lambda_1,.....,\lambda_M\}; \ \det[\lambda_k \mathbf{I} - \mathbf{M}] = 0;$$

$$\mathbf{M} = \mathbf{UGU^{-1}}; \ \ \mathbf{G} = \begin{vmatrix} \lambda_1 & 0 & \\ 0 & ... & \\ & & \lambda_M \end{vmatrix}; \ \ \mathbf{U} = [\mathbf{Y}_1, \mathbf{Y}_2, ...\mathbf{Y}_M]; \ \mathbf{M} \cdot \mathbf{Y_k} = \lambda_k \mathbf{Y_k}; \ k = 1,...M \qquad \text{(E-5)}$$

An arbitrary analytical matrix function of **M** can be expended into Taylor series and reduced to the function of its Jordan matrix **G** :

$$f(\mathbf{M}) = \sum_{i=1}^{\infty} f_i \mathbf{M}^i = \sum_{i=1}^{\infty} f_i \left(\mathbf{UGU^{-1}}\right)^i \equiv \left(\sum_{i=1}^{\infty} f_i \mathbf{U} (\mathbf{G})^i \mathbf{U^{-1}} \right) = \mathbf{U} \left(\sum_{i=1}^{\infty} f \ \ (\mathbf{G})^i \right) \mathbf{U}^{-1} = \mathbf{U} f(\mathbf{G}) \mathbf{U}^{-1} \quad \text{(E-6)}$$

Before embracing complicated things, let's look at trivial case, when Jordan matrix is diagonal:

$$f(\mathbf{G}) = \sum_{i=1}^{\infty} f_i \mathbf{G}^i = \sum_{i=1}^{\infty} f_i \begin{bmatrix} \lambda_1 & 0 & \\ 0 & ... & \\ & & \lambda_M \end{bmatrix}^i = \begin{bmatrix} \sum_{i=1}^{\infty} f_i \lambda_1^{\ i} & 0 & \\ 0 & ... & \\ & & \sum_{i=1}^{\infty} f_i \lambda_M^{\ i} \end{bmatrix} = \begin{bmatrix} f(\lambda_1) & 0 & \\ 0 & ... & \\ & & f(\lambda_M) \end{bmatrix} \text{(E-7)}$$

$$f(\mathbf{M}) = \mathbf{U} \begin{bmatrix} f(\lambda_1) & 0 & \\ 0 & ... & \\ & & f(\lambda_M) \end{bmatrix} \mathbf{U}^{-1}$$

The last expression can be rewritten as a sum of a product of matrix U containing only specific eigen vector (other columns are zero!) with matrix $\mathbf{U}^{-1}$:

$$f(\mathbf{M}) = [\mathbf{Y}_1 ... \mathbf{Y_k} ... \mathbf{Y}_M] \cdot \begin{vmatrix} f(\lambda_1) & 0 & \\ 0 & ... & \\ & & f(\lambda_M) \end{vmatrix} \mathbf{U^{-1}} = \sum_{k=1}^{M} f(\lambda_k) [0....\mathbf{Y_k}...0] \mathbf{U^{-1}} \qquad \text{(E-8)}$$

Still both eigen vector and $\mathbf{U}^{-1}$ in is very complicated (and generally unknown) functions of **M**…. Hmmmmm! We only need to find a matrix operator, which makes projection onto individual eigen vector. Because all eigen values are different, we have a very clever and simple way of designing projection operators. Operator

$$\mathbf{P}_k^i = \frac{\mathbf{M} - \lambda_k \mathbf{I}}{\lambda_i - \lambda_k} \qquad \text{(E-9)}$$

has two important properties: it is unit operator for $\mathbf{Y_i}$, it is zero operator for $\mathbf{Y_k}$ and multiply the rest of them by a constant:

$$
\begin{aligned}
&\mathbf{P}_k^i\mathbf{Y}_k = \frac{\mathbf{M}\cdot\mathbf{Y}_k - \lambda_k\mathbf{I}\cdot\mathbf{Y}_k}{\lambda_i-\lambda_k} = \frac{\lambda_k-\lambda_k}{\lambda_i-\lambda_k}\mathbf{Y}_k \equiv 0;\\
&\mathbf{P}_k^i\mathbf{Y}_i = \frac{\mathbf{M}\cdot\mathbf{Y}_i - \lambda_k\mathbf{I}\cdot\mathbf{Y}_i}{\lambda_i-\lambda_k} = \frac{\lambda_i-\lambda_k}{\lambda_i-\lambda_k}\mathbf{Y}_i \equiv \mathbf{Y}_i;\\
&\mathbf{P}_k^i\mathbf{Y}_j = \frac{\mathbf{M}\cdot\mathbf{Y}_j - \lambda_k\mathbf{I}\cdot\mathbf{Y}_j}{\lambda_i-\lambda_k} = \frac{\lambda_j-\lambda_k}{\lambda_i-\lambda_k}\mathbf{Y}_j
\end{aligned}
\tag{E-10}
$$

I.e. it project U into a subspace orthogonal to $\mathbf{Y_k}$. We should note the most important quality of this operator: it comprises of known matrixes: **M** and unit one. Also, zero operators for two eigen vectors commute with each other – being combination of **M** and **I** makes it obvious. Constructing unit projection operator $\mathbf{Y_i}$ which is also zero for remaining eigen vectors is straight forward from here: it is a product of all M-1 projection operators

$$
\begin{aligned}
&\mathbf{P}_{unit}^i = \prod_{k\neq i}\mathbf{P}_k^i = \prod_{k\neq i}\left(\frac{\mathbf{M}-\lambda_k\mathbf{I}}{\lambda_i-\lambda_k}\right);\\
&\mathbf{P}_{unit}^i\mathbf{Y}_j = \delta_j^i\mathbf{Y}_j = \left\{\begin{array}{c}\mathbf{Y}_i,\, j=i\\ \mathbf{O},\, j\neq i\end{array}\right\}
\end{aligned}
\tag{E-11}
$$

Observation that

$$
\mathbf{P}_{unit}^k\mathbf{U} = \mathbf{P}_{unit}^k[\mathbf{Y}_1...\mathbf{Y_k}...\mathbf{Y}_M] = [0....\mathbf{Y_k}...0] \tag{E-12}
$$

allows us to rewrite eq. (E-8) in the form which is easy to use:

$$
f(\mathbf{M}) = \sum_{k=1}^{M} f(\lambda_k)[0....\mathbf{Y_k}...0]\mathbf{U^{-1}} = \sum_{k=1}^{M} f(\lambda_k)\mathbf{P}_{unit}^k\mathbf{U}\cdot\mathbf{U^{-1}} = \sum_{k=1}^{M} f(\lambda_k)\mathbf{P}_{unit}^k; \tag{E-13}
$$

which with (E-11) give final form of Sylvester formula (E-for non-degenerated matrixes):

$$
f(\mathbf{M}) = \sum_{k=1}^{M} f(\lambda_k)\prod_{i\neq k}\left(\frac{\mathbf{M}-\lambda_i\mathbf{I}}{\lambda_k-\lambda_i}\right); \tag{E-14}
$$

One can see that this is a polynomial of power M-1 of matrix M, as we expected from the theorem of Jordan and Kelly that matrix is a root of its characteristic equation:

$$
g(\lambda) = \det[\mathbf{M}-\lambda\mathbf{I}];\; g(\mathbf{M}) \equiv 0; \tag{E-15}
$$

which is polynomial of power M. It means that any polynomial of higher order of matrix **M** can reduced to M-1 order. Equation (E-14) gives specific answer how it can be done for the arbitrary series.

If matrix **M** is reducible to diagonal form, where some eigen values have multiplicity, we need to sum only by independent eigen values:

$$
f(\mathbf{M}) = \sum_{k=1}^{m} f(\lambda_k)\prod_{\lambda_i\neq\lambda_k}\left(\frac{\mathbf{M}-\lambda_i\mathbf{I}}{\lambda_k-\lambda_i}\right); \tag{E-14-red}
$$

and it has maximum power of **M** of m-1. Prove it trivial using the above.

Let's return to most general case of Jordan blocks, i.e. a degenerated case when eigen values have non-unit multiplicity. For a general form of the Jordan matrix we can only say that it is direct sum of the function of the Jordan blocks:

$$f(\mathbf{G}) = \sum_{i=0}^{\infty} f_i \mathbf{G}^i = \sum_{i=0}^{\infty} f_i \begin{bmatrix} \mathbf{G_1^1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{...} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{...} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & \mathbf{G_m^{p_m}} \end{bmatrix}^i = \begin{bmatrix} \sum_{i=0}^{\infty} f_i \left(\mathbf{G_1^1}\right)^i & 0 & \\ 0 & ... & \\ & & \sum_{i=0}^{\infty} f_i \left(\mathbf{G_m^{p_m}}\right)^i \end{bmatrix}$$

$$= \begin{bmatrix} f(\mathbf{G_1^1}) & 0 & \\ 0 & ... & \\ & & f(\mathbf{G_m^{p_m}}) \end{bmatrix} == \sum_{\oplus k=1,m,\ \ j=1,p_k} f\left(\mathbf{G_k^j}\right) = f\left(\mathbf{G_1^1}\right) \oplus \mathbf{....} \oplus f\left(\mathbf{G_m^{p_m}}\right); \qquad \text{(E-16)}$$

Function of a Jordan block of size n contains not only the function of corresponding eigen value λ, but also its derivatives to (n-1)th order:

$$\mathbf{G} = \begin{bmatrix} \lambda & 1 & ... & 0 \\ 0 & \lambda & ... & 0 \\ ... & ... & ... & ... \\ 0 & 0 & ... & 1 \\ 0 & 0 & ... & \lambda \end{bmatrix}; f(\mathbf{G}) = \begin{bmatrix} f(\lambda) & f'(\lambda)/1! & ...f^{(k)}(\lambda)/k! & f^{(n-1)}(\lambda)/(n-1)! \\ 0 & f(\lambda) & ... & f^{(n-2)}(\lambda)/(n-2)! \\ ... & .... & ... & ... \\ 0 & 0 & ... & f'(\lambda)/1! \\ 0 & 0 & ... & f(\lambda) \end{bmatrix} \qquad \text{(E-17)}$$

The prove is attached in Appendix Eq. 17. We are half-way through.

There is sub-space of eigen vectors $\mathcal{Y}_k^n$ which corresponds to to the eigen value λ_k and the block $\mathbf{G_k^n}$:

$$\mathcal{Y}_k^n \in \left\{\mathbf{Y_k^{n,1}},....,\mathbf{Y_k^{n,q}}\right\}; \quad q = size\left(\mathbf{G_k^n}\right) \qquad \text{(E-18)}$$

$$\mathbf{M} \cdot \mathbf{Y_k^{n,1}} = \lambda_k \mathbf{Y_k^{n,1}}; \quad \mathbf{M} \cdot \mathbf{Y_k^{n,\mathit{l}}} = \lambda_k \mathbf{Y_k^{n,\mathit{l}}} + \mathbf{Y_k^{n,\mathit{l}-1}}; \ 1 < l \le q \qquad \text{(E-19)}$$

It is obvious from equation (E-17) that projection operator (E-11) will not be zero operator for $\mathcal{Y}_k^n$, and it also will not be unit operator for $\mathcal{Y}_i^n$. Now, let's look on how we can project on individual sub-spaces, eigen vectors, including zero-operator for specific sub-spaces. Just step by step (from eq. (E-6) and (E-17):

$$f(\mathbf{M}) = \mathbf{U} f(\mathbf{G}) \mathbf{U}^{-1}$$

$$\mathbf{U} f(\mathbf{G}) = \sum_{k=1}^{m} \sum_{i=1}^{n_k-1} \frac{f^{(i)}(\lambda_k)}{i!} \begin{bmatrix} \underbrace{0}_{\lambda_1} & \underbrace{0}_{\lambda_2} & ... & \underbrace{0}_{\lambda_{k-1}} & \underbrace{A_k^i}_{\lambda_k} & ...0... & \underbrace{0}_{\lambda_m} \end{bmatrix} \qquad \text{(E-20)}$$

$$A_k^i = \begin{bmatrix} \underbrace{B_1^{i\ k} \quad .... \quad B_{p_k}^{i\ k}}_{\lambda_k} \end{bmatrix}; B_n^{i\ k} = \begin{bmatrix} \underbrace{0.....0}_{i \ \ collumns} & Y_k^{n,1} & Y_k^{n,q_n-1} \end{bmatrix} \qquad \text{(E-21)}$$

i.e.

$$\mathbf{U}f(\mathbf{G}) = \sum_{k=1}^{m}\sum_{i=1}^{n_k-1}\frac{f^{(i)}(\lambda_k)}{i!}\left[\underbrace{0}_{\lambda_1}\quad\underbrace{0}_{\lambda_2}\quad\cdots\quad\underbrace{0}_{\lambda_{k-1}}\quad\underbrace{\left[\underbrace{\cdots}_{1}\quad\cdots\quad\underbrace{\underbrace{0.....0}_{i\ collumns}\quad Y_k^{n,1}\quad Y_k^{n,q_n-1}}_{n-th}\quad\underbrace{\cdots}_{p_k}\right]}_{\lambda_k}\quad ...0...\quad\underbrace{0}_{\lambda_m}\right] \tag{E-22}$$

From (E-19) we get:

$$\begin{aligned}
&[\mathbf{M}-\lambda_k\mathbf{I}]\cdot\mathbf{Y}_{\mathbf{k}}^{\mathbf{k},q} = 0;\quad [\mathbf{M}-\lambda_k\mathbf{I}]\cdot\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},k} = \mathbf{Y}_{\mathbf{k}}^{\mathbf{n},k-1};\ \ 1<k\le q\\
&U_1^{\mathbf{n}\ \ \mathbf{k}} = \left[\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},1}...\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},l}...\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},q}\right];\\
&[\mathbf{M}-\lambda_k\mathbf{I}]\cdot U_1^{\mathbf{n}\ \ \mathbf{k}} = U_2^{\mathbf{n}\ \ \mathbf{k}} = \left[0,\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},1}...\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},l}...\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},q-1}\right]\\
&......\\
&[\mathbf{M}-\lambda_k\mathbf{I}]^j\cdot U_1^{\mathbf{n}\ \ \mathbf{k}} = U_j^{\mathbf{n}\ \ \mathbf{k}} = \left[\underbrace{0..0}_{j\ \ zeros},\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},1}...\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},l}...\mathbf{Y}_{\mathbf{k}}^{\mathbf{n},q-j}\right]\\
&.....\\
&[\mathbf{M}-\lambda_k\mathbf{I}]^q\cdot U_1^{\mathbf{n}\ \ \mathbf{k}} = 0
\end{aligned} \tag{E-23}$$

$$\mathbf{U}f(\mathbf{G}) = \sum_{k=1}^{m}\sum_{i=1}^{n_k-1}\frac{f^{(i)}(\lambda_k)}{i!}[\mathbf{M}-\lambda_k\mathbf{I}]^i\left|\underbrace{0}_{\lambda_1}\quad\underbrace{0}_{\lambda_2}\quad\cdots\quad\underbrace{0}_{\lambda_{k-1}}\quad\underbrace{\left[U^{k\ \ 1}\quad\cdots\quad\underbrace{U^{k\ \ n}}_{n-th}\quad U^{k\ \ p_k}\right]}_{\lambda_k}\quad ...0...\quad\underbrace{0}_{\lambda_m}\right| \tag{E-24}$$

i.e. we collected all eigen vectors belonging to the eigen value λ_k. Now we need a projection non-distorting operator on the sub-space of λ_k. First, let's find zero operator for sunspace of λ_i:

$$\begin{aligned}
&O_i = [\mathbf{M}-\lambda_i\mathbf{I}]^{n_i} \Rightarrow [\mathbf{M}-\lambda_i\mathbf{I}]^{n_i}\,U_1^{r\ \ i} = [\mathbf{M}-\lambda_i\mathbf{I}]^{n_i}\left[\mathbf{Y}_{\mathbf{k}}^{r,1}...\mathbf{Y}_{\mathbf{k}}^{r,l}...\mathbf{Y}_{\mathbf{k}}^{r,q}\right] = 0;\\
&T_k = \prod_{i\ne k}\frac{O_i}{(\lambda_k-\lambda_i)^{n_i}} = \prod_{i\ne k}\left(\frac{\mathbf{M}-\lambda_i\mathbf{I}}{\lambda_k-\lambda_i}\right)^{n_i}
\end{aligned} \tag{E-25}$$

T_k is projection operator of sub-space of λ_k, but it is not unit one! To correct that we need an operator which we crate as follows:

$$\begin{aligned}
&R = \frac{\mathbf{M}-\lambda_i\mathbf{I}}{\lambda_k-\lambda_i};\quad T = \mathbf{M}-\lambda_k\mathbf{I};\quad \alpha = \alpha_{k,i} = 1/(\lambda_k-\lambda_i)\\
&RU_1 = U_1 + \alpha U_2 \qquad\qquad U_1 = U_1\\
&.....\\
&RU_{q-1} = U_{q-1} + \alpha U_q \qquad U_{q-1} = T^{q-2}U_1\\
&RU_q = U_q \qquad\qquad\qquad U_q = T^{q-1}U_1
\end{aligned}$$

$$Q = \alpha T$$
$$U_q = RU_q = RT^{q-1}U_1$$
$$U_{q-1} = R(I+Q)U_{q-1} = RQT^{q-2}U_1$$
$$U_{q-1} = RQU_{q-1} = RQT^{q-2}U_1$$
.....
$$U_1 = R\left(\sum_{j}^{q-1} Q^j\right)U_1$$

so, we get it:

$$P_k^i = \frac{\mathbf{M}-\lambda_i\mathbf{I}}{\lambda_k-\lambda_i}\left(\mathbf{I}+\sum_{j=1}^{n_k-1}\left(\frac{\mathbf{M}-\lambda_k\mathbf{I}}{\lambda_i-\lambda_k}\right)^j\right) \tag{E-26}$$

The final stroke is:

$$P_k = \prod_{i\neq k}\left(P_k^i\right)^{n_i} = \prod_{i\neq k}\left\{\frac{\mathbf{M}-\lambda_i\mathbf{I}}{\lambda_k-\lambda_i}\left(\mathbf{I}+\sum_{j=1}^{n_k-1}\left(\frac{\mathbf{M}-\lambda_k\mathbf{I}}{\lambda_i-\lambda_k}\right)^j\right)\right\}^{n_i} \tag{E-27}$$

and

$$F(M) = \sum_{k=1}^{m}\left[\prod_{i\neq k}\left\{\frac{\mathbf{M}-\lambda_i\mathbf{I}}{\lambda_k-\lambda_i}\left(\mathbf{I}+\sum_{j=1}^{n_k-1}\left(\frac{\mathbf{M}-\lambda_k\mathbf{I}}{\lambda_i-\lambda_k}\right)^j\right)\right\}^{n_i}\sum_{i=1}^{n_k-1}\frac{f^{(i)}(\lambda_k)}{i!}\left[\mathbf{M}-\lambda_k\mathbf{I}\right]^i\right] \tag{E-28}$$

Proof of eq. (E-17):

$$\mathbf{G}^0 = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & \ldots & 0 \\ 0 & 0 & & 1 \end{bmatrix}; \quad \mathbf{G}^1 = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \ldots & 1 \\ 0 & 0 & & \lambda \end{bmatrix}$$

$$\mathbf{G}^2 = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \ldots & 1 \\ 0 & 0 & & \lambda \end{bmatrix}\cdot\begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \ldots & 1 \\ 0 & 0 & & \lambda \end{bmatrix} = \begin{bmatrix} \lambda^2 & \lambda & 1\ldots & 0 \\ 0 & \lambda^2 & \lambda\ldots & 0 \\ 0 & 0 & \ldots & \lambda \\ 0 & 0 & 0 & \lambda^2 \end{bmatrix}$$

Induction:

$$\mathbf{G}^n == \begin{bmatrix} \lambda^n & n\lambda^{n-1}/1! & n(n-1)\lambda^{n-1}/2! & \dots & \dots \\ 0 & \lambda^n & n\lambda^{n-1}/1! & & \dots \\ 0 & 0 & \dots & & n\lambda^{n-1}/1! \\ 0 & 0 & 0 & & \lambda^n \end{bmatrix}$$

$$\mathbf{G}^2 = \begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \dots & 1 \\ 0 & 0 & & \lambda \end{bmatrix} \cdot \begin{bmatrix} \lambda^n & n\lambda^{n-1}/1! & n(n-1)\lambda^{n-1}/2! & \dots & \dots \\ 0 & \lambda^n & n\lambda^{n-1}/1! & & \dots \\ 0 & 0 & \dots & & n\lambda^{n-1}/1! \\ 0 & 0 & 0 & & \lambda^n \end{bmatrix} =$$

$$\begin{bmatrix} \lambda^{n+1} & (n+1)\lambda^n/1! & (n(n-1)+2n)\lambda^{n-1}/2! & \dots & \dots \\ 0 & \lambda^{n+1} & (n+1)\lambda^n/1! & & \dots \\ 0 & 0 & \dots & & (n+1)\lambda^n/1! \\ 0 & 0 & 0 & & \lambda^{n+1} \end{bmatrix}$$

$$\begin{bmatrix} \lambda^n & C_1^n\lambda^{n-1} & C_2^n\lambda^{n-2} & \dots & C_k^n\lambda^{n-k} & C_{k+1}^n\lambda^{n-k-1} & .. \\ 0 & \lambda^n & C_1^n\lambda^{n-1} & \dots & C_{k-1}^n\lambda^{n+1-k} & C_k^n\lambda^{n-k} & \\ .... & .... & .... & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda^n \end{bmatrix}$$

$$\begin{bmatrix} \lambda & 1 & 0 & 0 \\ 0 & \lambda & 1 & 0 \\ 0 & 0 & \dots & 1 \\ 0 & 0 & & \lambda \end{bmatrix} \cdot \begin{bmatrix} \lambda^n & C_1^n\lambda^{n-1} & C_2^n\lambda^{n-2} & \dots & C_k^n\lambda^{n-k} & C_{k+1}^n\lambda^{n-k-1} & .. \\ 0 & \lambda^n & C_1^n\lambda^{n-1} & \dots & C_{k-1}^n\lambda^{n+1-k} & C_k^n\lambda^{n-k} & \\ .... & .... & .... & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda^n \end{bmatrix} =$$

$$\begin{bmatrix} \lambda^{n+1} & (C_1^n+1)\lambda^n & (C_2^n+C_1^n)\lambda^{n-2} & \dots & (C_k^n+C_{k-1}^n)\lambda^{n-k+1} & (C_{k+1}^n+C_k^n)\lambda^{n-k} & .. \\ 0 & \lambda^{n+1} & (C_1^n+1)\lambda^n & \dots & .... & (C_k^n+C_{k-1}^n)\lambda^{n-k+1} & \\ .... & .... & .... & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & 0 & 0 & 0 & \lambda^{n+1} \end{bmatrix}$$

polynomial coefficients: $C_k^{n+1} = C_k^n + C_{k-1}^n;\ C_k^n = n!/k!/(n-k)!$ proves the point.

Hence, we can now calculate a polynomial functions or any function expandable into a Taylor series:

$$f(\mathbf{G})=\sum_{n=0}^{\infty} f_n \mathbf{G}^n=\sum_{n=0}^{\infty} f_n \begin{bmatrix} \lambda^n & C_1^n \lambda^{n-1} & \ldots . C_k^n \lambda^{n-k} \ldots & \ldots \\ 0 & & \ldots & \ldots . \\ 0 & & \ldots & \lambda^n \end{bmatrix}^i = \begin{bmatrix} \sum_{n=0}^{\infty} f_n \lambda^n \ldots & \sum_{n=0}^{\infty} f_n C_k^n \lambda^{n-k} & \ldots . \\ 0 & \ldots & \ldots . \\ 0 & 0 & \sum_{n=0}^{\infty} f_n \lambda^n \end{bmatrix}$$

The final stroke is noting that

$$\sum_{n=0}^{\infty} f_n C_k^n \lambda^{n-k}=\sum_{n=0}^{\infty} f_n \cdot \frac{n!\lambda^{n-k}}{k!\cdot(n-k)!}=\frac{1}{k!}\sum_{n=0}^{\infty} f_n \cdot \frac{n!\lambda^{n-k}}{(n-k)!}=\frac{1}{k!}\sum_{n=0}^{\infty} f_n \cdot \lambda^{n-k}\prod_{j=0}^{k-1}(n-j)$$

$$=\frac{1}{k!}\frac{d^k}{d\lambda^k}\sum_{n=0}^{\infty} f_n \cdot \lambda^n=\frac{1}{k!}\frac{d^k f}{d\lambda^k} \quad \#$$

Good HW exercise.