BINOMIAL THEOREM APPLICATIONS IN MATRIX FRACTIONAL POWERS CALCULATION

A. ARIAS, E. GUTIÉRREZ and E. POZO

Universidad del Pais Vasco, Escuela Superior de Ingenieros de Bilbao, Spain
Received Aug. 9, 1988

Fractional powers of matrix can be useful in many technical problems. Some authors think historical, economical, genetic future events may not be predicted.

However if we suppose some simplified hypothesis in the natural, social, phenomenous behaviours, the problems really changes. It actually changed since A. A. Markov (1856–1922) developed the stochastic process theory.

Let a Markov process or a Markov chain example. The Markov chains are based on the hypothesis that the events in a determinated step, only depend on the anterior step, and this remains all the time. We can express this dependence with a transition probabilities matrix. Future or pass situations can be predicted, when the matrix is well-known, and then the problem may be reduced to calculate fractional or natural powers of the matrix.

Other differents matrix applications exist, of course.

Communications matrices are really useful in sociology to get conclusions about the supremacy of certains groups of people over other groups. In this case, the problem is to calculate the periodicity order of the permutation matrix.

We can also apply the method developed in this paper to the linear analysis of structures and to the resolution of continuous medium problems by the finite elements theory, to calculate the vibrational frequency of structures. The iterative algorithms are not very suitable to solve linear problems. On the other hand, its efficacity is specially important in some non linear problems. This method is easy to implement and it could serve us to solve this kind of problems. The problem consists to compute A^p, (p integer of fractional).

The technique developed in this paper is based on the Binomial Theorem. The matrix will be expressed by the sum of the unit matrix and other one: $\mathbf{A} = \mathbf{I} + \mathbf{B}$.

Applying the Binomial Theorem, we will obtain a matrices powers series:

$$\mathbf{A} = (\alpha_{ij}); \; \mathbf{A}^p = (\mathbf{I} + \mathbf{B})^p.$$

Then, the problem is to verify the convergence properties of that series. The power series will converge only if the matrix norm is a very small real number: $||\mathbf{B}|| \ll 1$. This can be solve introducing an appropriate scalar k, $(\mathbf{A} = k(\mathbf{I} + \mathbf{M}))$, that will be optimally calculate in terms of the original matrix coefficients. That is, $||\mathbf{M}||_1$ is minimized making $\delta||\mathbf{M}||_1/\delta k = 0$.

Method

The Newton Theorem can be expressed by the next expression:

$$(1 + X)^p = 1 + pX + (P_2)\, X^2 + (P_3)X^3 + \ldots \tag{1}$$

being $|X| < 1 \Leftrightarrow X^2 < 1$
$p \in$ integer or fractional numbers.

The sum of the first several terms of (1) can be used as an approximation to $(1 + X)^p$.

When the elements in (1) are matrices the expression is:

$$\mathbf{A}^p = (\mathbf{I} + \mathbf{B})^p = \mathbf{I} + p\mathbf{B} + (p(p-1)/2\,!)\, \mathbf{B}^2 + \ldots + (P_n)B^n + \ldots$$

The series $a_1 + a_2 + \ldots + a_n + \ldots$ converges if:

$$\lim_{n\to\infty} ||a_{n+1}/a_n|| < 1\,.$$

In our case:

$$||a_{n+1}/a_n||^{2n} = ||(P_{n+1})\,\mathbf{B}^{n+1}/(P_n)\,\mathbf{B}^n|| = |p(p-1)\,\ldots$$

$$\ldots (p-n)n\,!/(1+n)\,!p(p-1)\ldots(p-n+1)|\ \ ||\mathbf{B}|| =$$

$$= |p - n/(n+1)|\ \ ||\mathbf{B}||;$$

$$\lim_{n\to\infty} |p - n/(n+1)|\ \ ||\mathbf{B}|| = ||\mathbf{B}||\,.$$

Then, the series (2) converges if $||\mathbf{B}|| < 1$.
Remembering the habitual norms of a matrix:

$$||\mathbf{B}||_1 = (\sum_i \sum_j b_{ij}^2)^{1/2}$$

$$||\mathbf{B}||_2 = (\max_i \sum_j |b_{ij}|.$$

Remembering the property:

$$||\mathbf{B}^p|| \leq ||\mathbf{B}||^p \text{ then:}$$

$$||\lim_{p\to\infty} \mathbf{B}^p|| = \lim_{p\to\infty} ||\mathbf{B}^p|| \leq$$

$$\leq \lim_{p\to\infty} ||\mathbf{B}||^p = 0 \Leftrightarrow \lim_{p\to\infty} \mathbf{B}^p = 0.$$

We could prove the convergence too, computing the characteristic roots of the matrix, if the highs of them is less than the unit. But, usually, the calculation of the eigen values is difficult.

Then $B^p \approx 0$ when $p \to \infty$.

To accelerate the convergence, we can do:

$$\mathbf{A} = k(\mathbf{I} + \mathbf{C}) \text{ where } \mathbf{C} = (1/k)\mathbf{A} - \mathbf{I} \text{ (being } k \text{ a scalar).}$$

Therefore $\mathbf{A}^p = k^p(\mathbf{I} + \mathbf{C})P = K^p(\mathbf{I} + p\mathbf{C} + (P_2)\mathbf{C}^2 + \ldots + (P_n)\mathbf{C}^n + \ldots)$.

We have to choose a value of k making $||\mathbf{C}|| < 1$, and as small as possible.

$$||\mathbf{C}||_1 = (\sum_i \sum_j \mathbf{C}_{ij}^2)1/2 \text{ and}$$

$$\text{if } i \neq j \;\; \mathbf{C}_{ij}^2 = 1/k^2 a_{ij}^2$$

$$\text{if } i = j \;\; \mathbf{C}_{ii}^2 = (1/k \cdot a_{ii} - 1)^2 = 1/k^2 a_{ii}^2 - 2/k a_{ii} + 1.$$

Then

$$||\mathbf{C}||_1 = ((1/k)^2 \sum_i \sum_j a_{ij}^2 - 2/k \sum_i a_{ii} + n)^{1/2}.$$

Minimizing $||\mathbf{C}||_1$

$$\delta||\mathbf{C}||_1/\delta k = 0 \Leftrightarrow$$

$$\delta||\mathbf{C}||_1/\delta k = (-2/k^3 \sum_i \sum_j a_{ij}^2 + 2/k^2 \sum_i a_{ii})/2||\mathbf{C}||_1 = 0$$

$$\Leftrightarrow (-1/k \sum_i \sum_j a_{ij}^2 + \sum_i a_{ii}) = 0 \Leftrightarrow$$

$$k = \sum_i \sum_j a_{ij}^2 / \sum_i a_{ii}.$$

A numerical application

Let $\mathbf{A}$ a transition matrix. $\mathbf{A} = \begin{bmatrix} 0.75 & 0.15 & 0.10 \\ 0.15 & 0.70 & 0.15 \\ 0.20 & 0.20 & 0.60 \end{bmatrix}$.

Our problem is to compute the matrix power $A^{1/p}$ (with p a natural number).

Using the Taylor's series: (1)

$$\mathbf{A}^{1/p} = (\mathbf{I} + \mathbf{B})^{1/p} = \mathbf{I} + (1/p_1)\mathbf{B} + (1/p_2)\mathbf{B}^2 + \ldots$$

The series will converge if: $||\mathbf{B}|| < 1$.

We suppose that $p = 5$ and we will truncate the series (1) in the term of 6th degree (without taking this last one).

The approximation is:

$$\mathbf{A}_1{}^{1/5} = \begin{bmatrix} 0.93865906 & 0.03672479 & 0.02461615 \\ 0.03512933 & 0.92375835 & 0.04111232 \\ 0.05162550 & 0.05322096 & 0.89515354 \end{bmatrix}.$$

In order to evaluate the approximation precision obtained, we evaluate:

$$(\mathbf{A}_1{}^{1/5})^5 = \mathbf{A}_1 = \begin{bmatrix} .75033585 & .14981167 & .09985248 \\ .15007531 & .70083799 & .14908670 \\ .19930953 & .19904588 & .60164458 \end{bmatrix}.$$

The induced metric by the norm $||\,.\,||_1$ which makes the matrices $\mathbf{A}$ and $\mathbf{A}_1$ different is:

($\mathbf{A}_1$ is the matrix approximated by the series (1))

$$d_1 = ||\mathbf{A}_1||_1 - ||\mathbf{A}||_2 = 1.2540578 - 1.2529964 = 0.00106142\,.$$

Now, we want to accelerate the convergence of the method described above. We make: $\mathbf{A} = k(\mathbf{I} + \mathbf{M})$

So as $\quad \mathbf{A}^{1/p} = K^{1/p}(\mathbf{I} + \mathbf{M})^{1/p} = k^{1/p}(\mathbf{I} + (1/p_1)\mathbf{M} + (1/p_2)\mathbf{M}_2 + \ldots). \qquad (2)$

In order to optimize the convergence of the series (2), we calculate the value of K that minimizes the norm of M. (we will choose the norm $||\,.\,||_1$). The value of K can be expressed in terms of the matrix components. $\mathbf{A} = (\alpha_{ij})$

$$k = \sum_i \sum_j \alpha_{ij}^2 \,/ \sum_i \alpha_{ii}\,.$$

In this case: $k = 1.57/2.05 = 0.76585366$.
Being $K^{1/5} = 0.94804545$.
So that, $\mathbf{M}$ will be:

$$\mathbf{M} = \begin{bmatrix} -0.02070064 & 0.19585987 & 0.13057325 \\ 0.19585987 & -0.08598726 & 0.19585987 \\ 0.26114650 & 0.26114650 & -0.21656051 \end{bmatrix}.$$

Being $||\mathbf{M} = ||_1 = 0.56854939$.

Developing $\mathbf{A}^{1/p}$ at series (2) and truncating in the same way, we obtain the approximation:

$$\mathbf{A}_2^{1/5} = \begin{bmatrix} .93856319 & .03678552 & .02466390 \\ .03510280 & .92349685 & .04141300 \\ .05185189 & .05353462 & .89462612 \end{bmatrix}.$$

The metric between $\mathbf{A}$ and $\mathbf{A}_2$ will be:

$$d_2 = ||\mathbf{A}_2||_1 - ||\mathbf{A}||_1 = 1.253149 - 1.2529964 = 0.00015265\,.$$

Conclusions

In many technical problems we need to evaluate fractional powers of a matrix. The Binomial Theorem is a really good tool to this aim. The difficulty appears when the norm of the matrix $||\mathbf{B}|| = ||\mathbf{A} - \mathbf{I}||$ is not less than the unit, and it is not small enough to assure the effective converge of the series used, either. Evaluating the scalar K, we will minimize the norm of the matrix $|\mathbf{M}||$ we have got, optimizing the series convergence.

$$||\mathbf{M}|| = |1/K| \, ||\mathbf{A} - \mathbf{I}|| \, .$$

Using the optimized method and truncating the series (2) at the same step, we have obtained an approximation of the solution that is 10 times closer than the solution that was obtained by the Binomial Theorem.

A. Arias
E. Gutiérrez
E. Pozo

Alda Urquijo S/N.
48013—Bilbao
Spain