A-posteriori error estimation for Nystrom method

Yuchen Su

December 20, 2017

Introduction

For this project, we explore a-posteriori error estimation for the numerical solution of Fredholm equations of the second kind using Nystrom's method. The ultimate goal is to develop an adaptive quadrature algorithm for the QBX quadrature rule or more generally error estimation techniques for Nystrom method discretizations. For a detailed review of Nystrom's method we refer the reader to Section 12.2 in [2]. In general the solution of a second kind equation

$$\varphi - A\varphi = f$$

is approximated using Nystrom's method by the equation

$$\varphi_n - A_n\varphi_n = f_n$$

we reduces to solving a finite linear systems. We derive an error estimator for the approximation $\varphi - \varphi_n$ by taking differences of the two equations:

$$\varphi - A\varphi - (\varphi_n - A_n\varphi_n) = f - f_n$$

Shifting the terms that include the operator to the right hand side and adding and subtracting $A_n\varphi$

$$\begin{aligned} \varphi\text{-}\varphi_n &= A\varphi - A_n\varphi_n + (A_n\varphi - A_n\varphi) + f - f_n \\ &= A\varphi - A_n\varphi + (A_n\varphi - A_n\varphi_n) + f - f_n \end{aligned}$$

Taking the term in the parantheses to the left handside, factor out $\varphi - \varphi_n$ we get

$$\begin{aligned} (I - A_n)(\varphi - \varphi_n) &= A\varphi - A_n\varphi + f - f_n \\ \Rightarrow \varphi - \varphi_n &= (I - A)_n^{-1}\left[\left(\left(A - A_n\right)\varphi\right) + \left(f - f_n\right)\right] \end{aligned} \quad (1)$$

We see from the last equation that the error in the density approximation by Nystrom is a function of $(A - A_n)\,\varphi$ which is in essence quadrature error and $f - f_n$ which is interpolation error of the function. Generally f is known before hand and so the interpolation error can be completely reduced for most purposes. The above derivation suggests that an error estimate for the quadrature rule can be used directly for the error estimate of **Nystroms** method. We explore this idea in the next section using some numerical examples.

Error Estimation for Quadrature Rules

There have been many linear error estimates used in adaptive quadrature rules an they can all be classified more or less into the following types (taken from [1]):

1. $\varepsilon \sim |Q_n^{(m_1)}[a,b] - Q_n^{(m_2)}[a,b]|$
2. $\varepsilon \sim |Q_{n_1}[a,b] - Q_{n_2}[a,b]|$

3. $\varepsilon \sim |f^{(n)}(\xi)|$
4. $\varepsilon \sim |\tilde{c}_n|$

Where $Q_n^{(m)}[a,b]$ denotes quadrature rule of the interval $[a,b]$ where the superscript (m) denotes the level of iteration in the adaptive quadrature (if necessary) and n denotes the number of quadrature points. The third estimate is based on approximating the derivative in the analystic error term and the fourth estimate is based on the highest-degree coefficient of truncated projection of the function onto some orthogonal basis. It has been shown in [3] that the first three error estimates are essentially the same so for ease, we focus on the second error estimate. Without looking at the analysis, it not hard to show numerically for some simple examples that these error estimates are not very good.

We consider the following one dimensional integral equations of the second kind.

Example 1 (1D Laplace kernel):

$$\varphi(x) - \frac{1}{2}\int_{-1}^{1} |x-y|\varphi(y)dy = g(x)$$

where

$$g(x) = e^x$$

We know the exact solution of the equation is

$$\varphi(x) = \frac{1}{2}xe^x + c_1e^x + c_2e^{-x}$$

where

$$\begin{aligned} c_1 &= c_2 + (e^2+1)^{-1} \\ c_2 &= \frac{e^4+6e^2+1}{8(e^2+1)} \end{aligned}$$

Example 2:

$$\varphi(x) - \frac{1}{2}\int_{-1}^{1} (x+1)e^{-xy}\varphi(y)dy = g(x)$$

where

$$g(x) = e^{-x} - \frac{1}{2}\left(e^{x+1} - x^{-(x+1)}\right)$$

This equation has the exact solution $\varphi = e^{-x}$.

We estimate the density of both these equations using Nystrom's method with the trapezoid quadrature rule and look at the error estimate given by $\varepsilon \sim |Q_{n_1}[a,b] - Q_{n_2}[a,b]|$.

Methodology

Since we don't have A exactly, but know that $A_n \to A$ pointwise (given a convergent quadrature rule ***Theorem 12.8*** [2]), we may reasonably choose to use both the operator A_{n_f} and density φ_{n_f} on a finely discretized grid as a proxy for A and φ. We choose $n_f = 100$ in which the density error $||\varepsilon_\varphi(n)|| = ||\varphi - \varphi_n||$

is for 0.0269 *Example 1* and 0.0093 for *Example 2.* In the following experiments, we fix a coarser grid on n quadrature points, compute the quadrature estimate of the integral at those n points and using Nystrom's method to interpolate the density approximation on the fine grid using thoses n quadrature points. The comparison is always made between the fine grid approximation using n quadrature points and the exact solution computed at the fine grid points (ie $\varphi_n, \varphi \in \mathbb{R}^{n_f}$).

Below we plot the density errors for the Trapezoid for for varying n. We see that the quadrature rule is a convergent one for both examples.

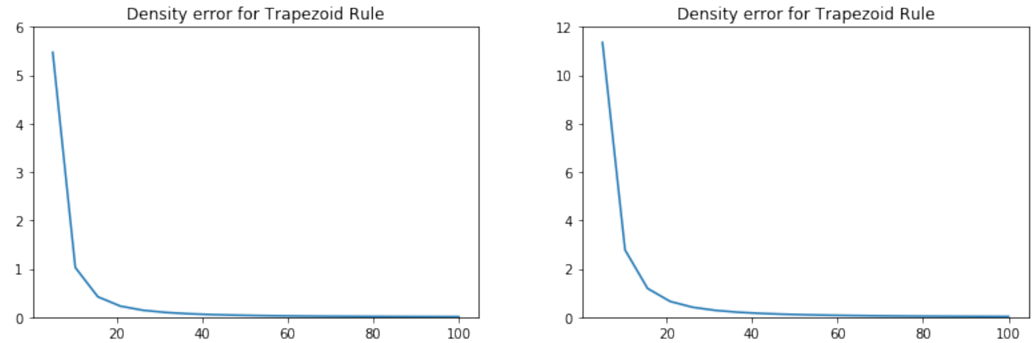

Figure 1: Density approximation errors of Trapezoid rule varying n (left: Example 1, right: Example 2)

Let's take a look at using the quadrature erorr estimate given by discretizations at n and $n+1$ quadrature points. First we plot the vector $\varepsilon_\varphi(n)$ for $n = 5, \ldots, 8$.

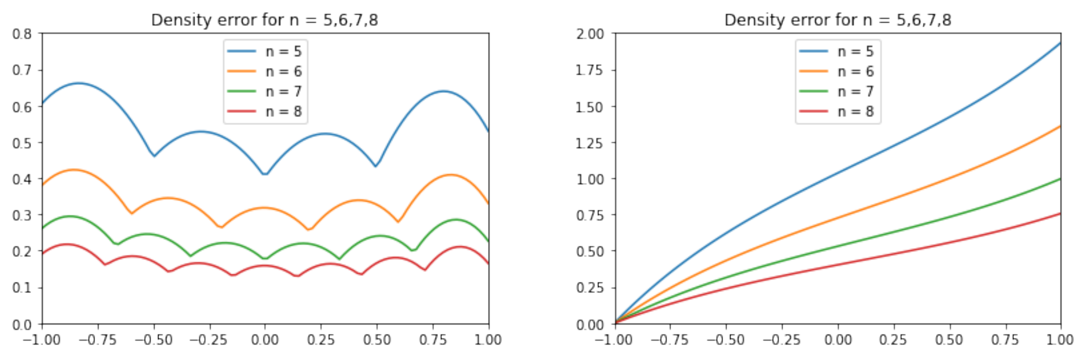

Figure 2: $\varepsilon_\varphi(n) = \varphi - \varphi_n$ for $n = 5, \ldots, 8$ (left: Example 1, right: Example 2)

We have essentially laid the error vector out flat. For example 1 we see that the density error is highest at the endpoints -1, and 1 and osscillates lower toward 0. For example 2 we see that the error increase monotonically from -1 to 1. The reason we plot the error vector and not the norm is that we want to see how the error behaves locally on our grid from $[-1, 1]$. A great error estimate can tell us exactly where in our mesh we need to refine or how much we need to refine a particular subset of our mesh relative to another portion to get the same accuracy. We plot below the error estimates for this above density error. First for Example 1

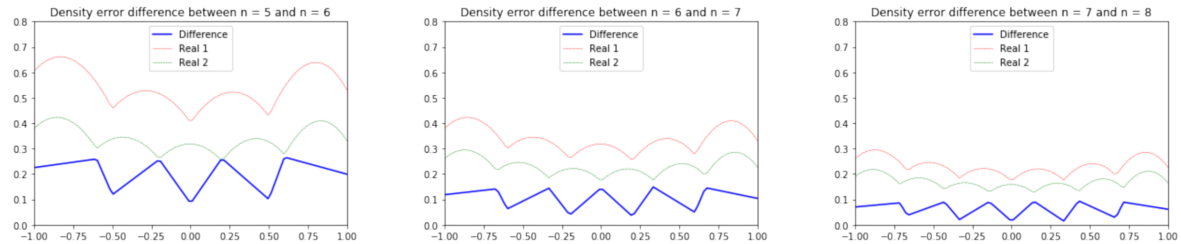

Figure 3: Quadrature Error Estimates for Example one using $Q_{n_1}[a, b] - Q_{n+1}[a, b]$

The actual errors are given in the red and green lines and the error estimate (the difference between the red and green lines) is given in the blue line. Consider using the blue line to estimate the error of the $n+1$ density error. The first thing one notices is that when you take norm, the error estimate is a lower bound on the error (which is the wrong direction!). Also it does not capture the fact that the error is greatest at the endpoints and decreases towards the center of the domain. Also notice that the peaks in the error estimate coincide with the nadirs of the actualy $n+1$ error. As an error estimate, we are not doing a good job! The graphs for example 2 tell a similar story:

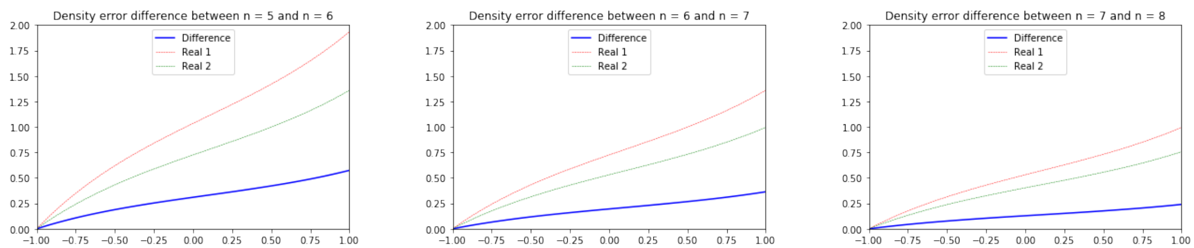

Figure 4: Quadrature Error Estimates for Example one using $Q_{n_1}[a, b] - Q_{n+1}[a, b]$

In all these cases, the error estimate grossly underestimates the error. Also the local information is greatly diluted. Based on these simple numerical tests, I would advise against putting too much faith in using quadrature error estimates directly to estimate Nystrom's error.

Anselone's Estimate

Can we do better than just using the Quadrature error estimate? Notice that in (1), the quadrature error is actually operated on by $(I - A_n)^{-1}$. The common theme in the previous section was that although the error estimates had captured the general theme, the local magnitudes were off. Perhaps this is because we ignored completely the action of $(I - A_n)^{-1}$ on the quadrature error. Luckily (or unluckily) for us, we can estimate this operator using the following

$$||(I - A_n)||^{-1} \leq \frac{1 + ||(I - A)^{-1} A_n||}{1 - ||(I - A)^{-1}(A_n - A)A_n||}$$

which is dude to Brakhage and Anselone and Moore and can be found as ***Theorem 10.12*** in **[2].** If we let $B_n = I + (I - A)^{-1} A_n$ and $S_n = (I - A)^{-1}(A_n - A)A_n$ and rearranging the above estimate (after removing the norms and a estimate via Neumann series) we get that

$$(I - S_n)(I - A_n)^{-1} = B_n$$

See page 191 in **[1]** for details. Note that by definition $||S_n|| \to 0$ as $n \to \infty$ and so as $n \to \infty$. As a result $I - S_n \to I$ and $B_n \to (I - A_n)^{-1}$ which is the quantity we want to estimate. In fact theres more. Since $A_n \to A$ pointwise, for any given problem B_n is actually approximating $(I - A)^{-1}$ as $n \to \infty$ however for our computational purposes, the prior interpretation is sufficient. So the idea is to take a look at B_n and see how the Nystrom discretation operates on the quadrature error. Perhaps this will fix the locality issues of the quadrature error estimate.

Computing Anselone's Estimate

Note that B_n has both a matrix inverse and a matrix multiplication so one must be very smart to approximate such operators if one is to use the operator computationally. We discuss some smart ways to do so in the next section, however for the purposes of just understanding what the creature B_n is, we brute force the actual computation of B_n.

The main problem in turning the operator B_n into something computable is that fact that A_n essentially is a finite dimensional object on n quadrature points and A is an infinite-dimensional operator. The trouble is finding the correct perspective to lift that corresponding spaces the operator A_n operators on to the infinite dimensional setting. Nystrom interpolation gives us a easy way to extend the range of A_n to arbitray dimensions, you simply compute the kernel for any x and so the dimension of the rows (range) of A_n can be extended to arbitrary dimensions. How then do we extend the domain (columns) of A_n to arbitary dimensions.

One way to reason about this is to consider the action of the operator $A_n\varphi$ on an infinite dimensional object. In the context of **Nystrom** the operator A_n only acts on the quantities of the density that live at the n quadrature points and so contributions elsewhere would not be taken into consideration. Essentially the dimensions outside of the select finite quadrature points are nullspace dimensions of the operator A_n. The easiest way to hack this behavior up computational is to insert 0 columns into the spaces of the density that correspond to dimensions outside of our quadrature points. Combining these two extensions, we may consider the operator A_n in the infinite dimensional setting and there comparison $A - A_n$ is well definte mathematically.

Below we plot some graphs of B_n. the methodology to compute this quantity as is follows. We take for a proxy of A, the discrete operator A_{n_f} on the fine grid of $n_f = 100$ points. For A_n, the columnspace of the operator lives at the n quadrature points and the row space is extended onto the fine grid by Nystrom interpolation. The columns are extended by putting zeroes in at the points of the fine grid in which do not coincide with the n quadrature points of the coarse grid. As a result, the grid used must be nested (which is another reason for the choice of a easier Newton-Coates method for these simple experiments). The inverese $(I - A_{n_f})^{-1}$ is computed directly as is the matrix-matrix product of the inverse with A_n. Essentially what is being computed is $B_n = I + (I - A_{n_f})^{-1}A_n$ where A_n is augmented to exist on the fine grid. We compare this quantity to the quantity B_n is estimating which is $B_R = (I - A_n)^{-1}$. We plot below the error matrix $|B_n - B_R|$ where $|\cdot|$ denotes an entry-wise absolutely value.

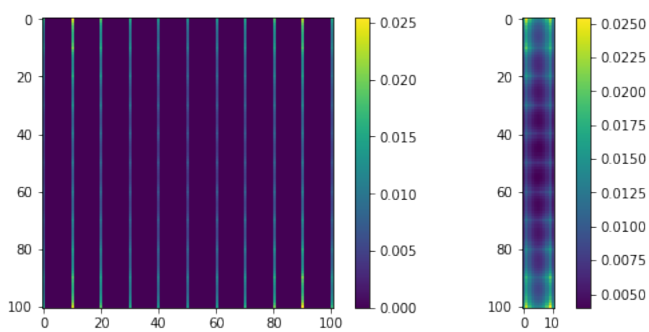

Figure 5: $|B_n - B_R|$ for Example 1

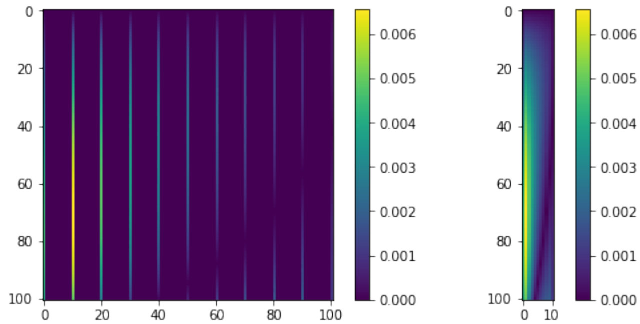

Figure 6: $|B_n - B_R|$ for Example 2

How does one intepret this? If I hypothesize that the matrix is capturing the local magnitudes of the error then I expect from looking at the actualy density error of the method then we should expect there to be error at the four corners of the matrix for Example 1 and near the later indices in both the rows and columns for Example 2. While these two error matrices do exhibit this somehwat it is not very convincing and also the magnitudes are all wrong! The error at the corners should be around 0.6 not 0.025 for Example 1.

It turns out however, this difference matrix is the wrong one to look at! It you think about it, we want the action of $(I - A_n)^{-1}$ and we get lost in the math and decided to compute B_n and look at the difference. Why not computed $(I - A_n)^{-1}$ directly? Even in our silly brute force computational world, computing this quantity makes us much happier. Well, what then do we want to compare this to? Well in the limit, $(I - A_n)^{-1} \to (I - A)^{-1}$ pointwise so why not use A on the fine grid. Below we plot the error matrix $|(I - A_n)^{-1} - (I - A_{n_f})^{-1}|$ for Example 1 and Example 2.

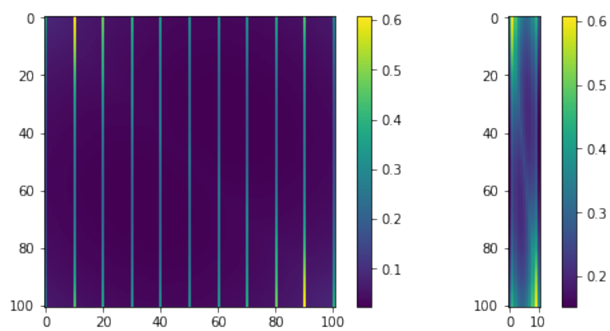

Figure 7: $|(I - A_n)^{-1} - (I - A_{n_f})^{-1}|$ for Example 1 ($n = 11$, $n_f = 101$)

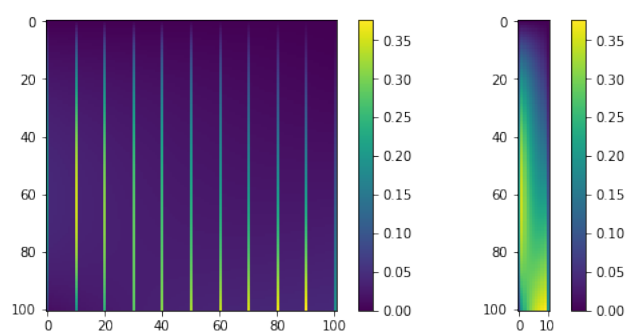

Figure 8: $|(I - A_n)^{-1} - (I - A_{n_f})^{-1}|$ for Example 2 ($n = 11$, $n_f = 101$)

This seems to look much better. The magnitudes are somewhat much more inline too (at least for Example

1).

Estimating Anselone

Whether or not the estimate $|(I - A_n)^{-1} - (I - A_{n_f})^{-1}|$ contains any information (I would argue it does), computationally a rational person would never consider using this as an actual error estimate. Therefore the next thing one asks is whether this quantity can be approximated somehow cheaply. The obvious estimator is $|(I - A_n)^{-1} - (I - A_{n_2})^{-1}|$ where $n < n_2 \ll n_f$. Below we plot these estimators:

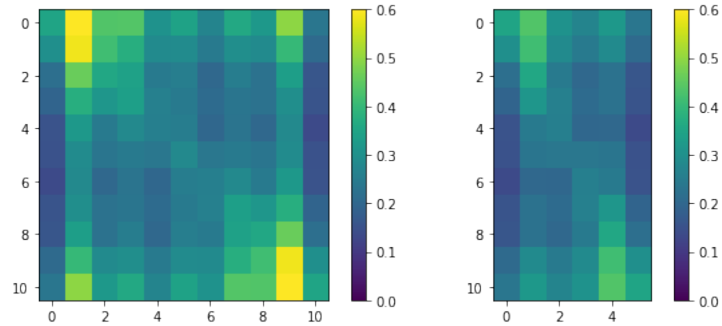

Figure 9: $|(I - A_n)^{-1} - (I - A_{n_f})^{-1}|$ for Example 1 ($n = 6$, $n_f = 11$)

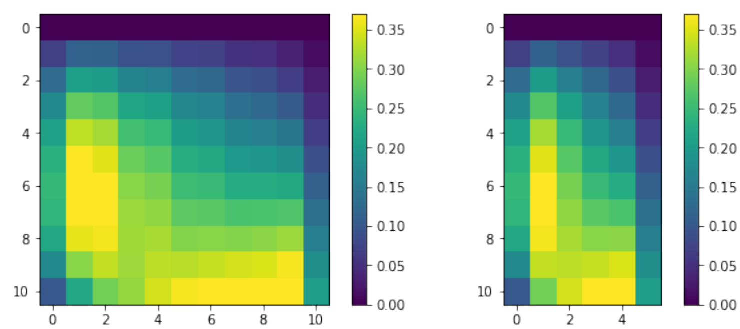

Figure 10: $|(I - A_n)^{-1} - (I - A_{n_f})^{-1}|$ for Example 2 ($n = 6$, $n_f = 11$)

This looks super promising. Not only is the local information of of the original error matrix $|(I - A_n)^{-1} - (I - A_{n_f})^{-1}|$ captured, but the magnitudes are as well!

A decomposed Nystrom error estimate & further computational speedups

What these experiments seem to suggests is that $|(I - A_n)^{-1} - (I - A_{n_2})^{-1}||Q_{n_1}[a, b] - Q_{n_2}[a, b]|$ should provide a reasonable Nystrom error estimate with the matrix error giving us the magnitude of the error and the quadrature error giving us some sort of shape (or maybe is is just not useful). The matrix error gives us some sort of local information. I can run some experiments after break to confirm this.

One may also ask whether it is even necessary to include $|Q_{n_1}[a, b] - Q_{n_2}[a, b]|$. If the information present in the operator is more important, we may be able to extract it more cheaply with a smart vector choice (rather than use the quadrature error vector). This forms a matrix-vector product estimator which in the context of PDEs, the kernels would allow fast computation using FMM.

The author has some ideas of how to do this, and will conduct some further experiments over break.

Further Work

We can done some very simple toy examples of a prospective method for error estimation for Nystrom method. The obvious line of work is to try things instead with Gauss-Legendre quadrature (this necessitates the search for non-smooth examples to play with). The matrix operators were depicted using nested grids and so we must consider Gauss-Kronrod quadrature along this route. The next thing to do after this is to look at integral equations that live on some complex boundary and finally to understand this method using a PDE and the QBX rule. Finally, we would like to make these things exact with analysis.

Conclusion

In conclusion, the quadrature error is a really bad error estimator for Nystrom method. It lacks local information on the magnitude of errors. We can perhaps salvage this using Anselone's estimate. We then realize that Anselone's estimate is a distraction and one can more easily (computationally) and accurately get the local magnitude information from looking at the error between $(I - A_n)^{-1}$ and $(I - A)^{-1}$. Well, I guess Anselone's estimate was not a distraction because ultimately reasoning that $B_n \to (I - A)^{-1}$ was kind of crucial to coming up with the thing that works better. Questions remain about how this cheaper extract this local information.

References

[1] Pedro Gonnet. A review of error estimation in adaptive quadrature. *ACM Computing Surveys (CSUR)*, 44(4):22, 2012.

[2] R. Kress. *Linear Integral Equations, 3rd ed.* Applied Mathematical Sciences. Springer New York, 2013.

[3] Dirk P. Laurie. Practical error estimation in numerical integration. *Journal of Computational and Applied Mathematics*, 12-13:20, 1985.