Extrapolation

Eesh Gupta

June 7, 2020

Extrapolation is a technique in which we deliberately make noise rate worse in order to get a more accurate result.

1 Background

1.1 Richardson Extrapolation

Suppose you have approximated $A^\star$ as $A(h)$ (assume h is the error rate) but you want to make a better approximation. Fortunately, you are also given a series

$$A^\star = A(h) + a_0 h^{k_0} + a_1 h^{k_1} + a_2 h^{k_2} + \dots \tag{1}$$

where a's are unknown constants and k's are known constants such that $h^{k_i} > h^{k_{i+1}}$. That is, the sequence on RHS of equation 1 is decreasing. If we write the same equation in terms of Big Oh notation, we have

$$A^\star = A(h) + O(h^{k_0}) \tag{2}$$

The current error is on the order of $O(h^{k_0})$. Our need is to get a better approximation of $A^\star$, let's say with a higher order error $O(h^{k_1})$ estimate. What do we do?

1. *Big Oh Notation* Rewrite the 1 as

   $$A^\star = A(h) + a_0 h^{k_0} + O(h^{k_1}) \tag{3}$$

   In the following steps, our objective is to get rid of the term $a_0 h^{k_0}$ so we end up with an improved error bound of $O(h^{k_1})$.

2. *Rescaling the Parameter* h as $\frac{h}{t}$ where t is a constant (hopefully $t \neq 1$!).

   $$A^\star = A(\frac{h}{t}) + a_0 \frac{h}{t}^{k_0} + O(h^{k_1}) \tag{4}$$

3. *Ridding of a term from equation 1:* Multiply 4 by t^{k_0}, subtract equation 3 from the modified 4 and divide both sides of the resulting equation by $t^{k_0} - 1$. Then we have

$$A^{\star} = \frac{t^{k_0} A(\frac{h}{t}) - A(h)}{t^{k_0} - 1} + O(h^{k_1}) \tag{5}$$

We can repeat this process many times for different error rates to improve the aprroximation further.

2 Mathematics of extrapolation

2.1 Series Expansion in Noise Parameter

Our starting point is the following eqution

$$\frac{\partial \rho}{\partial t} = -i[K(t), \rho] + \lambda \mathcal{L}(\rho) \tag{6}$$

where the first term on RHS corresponds the time dependent qubit hamiltonian (set of pauli gates acting on qubits with time dependent coupling coefficients) and the second term accounts for noise. Here, λ is the streingth of noise, assumed weak because of short depth circuits. While using a lindbladian superoperator may imply that we are considering only markovian noise, we can also account for non markovian noise by expanding $\mathcal{L}(\rho)$ as

$$\lambda \mathcal{L}(\rho) = -i[V, \rho] \tag{7}$$

where V is some hamiltonian.

Once we have a good starting point, we can then remove the time evolution due to qubit hamiltonian $K(t)$ to zoom in on the noisy evolution in the interation picture of $K(t)$. This involves “heisenburg-ing” our noisy superoperator $\mathcal{L}$ and “shrodinger-ing” our density matrix ρ to get the following partial differential equation:

$$\frac{\partial}{\partial t} \rho_I(t) = +\lambda \mathcal{L}_{I,t}(\rho_I(t)) \tag{8}$$

From here on, we use some tricks relating to perturbative expansion and go back to the schrodinger picture. After this process, we obtain the expectation value of some observale A as

$$E_K(\lambda) = E^{\star} + \sum_{k=1}^{n} a_k \lambda^k + R_{n+1}(\lambda, \mathcal{L}, T) \tag{9}$$

Here, $E_K(\lambda)$ is the expectation value of A given noise rate of λ, $E*$ is the noiseless expectation value of A, a_k hides huge integrals involving the noisy superoperator $\mathcal{L}$, R is the “remainder” of the infinite series and T is the time for which noisy evolution lasts.

Fortunately, R is bounded so the series is decreasing. Hence, we can apply richarson extrapolation methods to improve upon our approximation of noiseless expectation value $E^{\star}$.

2.2 Experimental Rescaling of the noise parameter

The chief concern in this section is how to control the noise parameters λ in an experimental setting. One possible solution is to rescale the time T for which the noisy evolution lasts. This may result in us either "waiting longer" to measure the results or "slowing down" the gate operations. Using some clever substitution techniques inside integrals, we can easily show that rescaling time T results in a rescaled noise rate λ.

2.3 Error bounds on the noise free parameter

Once we can rescale our noise rate λ, we are ready to perform Richardson Extrapolation. In our earlier discussion on this trick, we rescaled our noise just once to improve our error bound from $O(h^{k_0})$ to $O(h^{k_1})$. However, in general, we can perform multiple rescalings and improve our error bounds by multiple orders with only a few equations. So in this scenario, we will scale our noise rate n times and hence cancel n terms from RHS of equation 9. In doing so, we will obtain n expectation values

$$E_K^n(\lambda) = \sum_{j=0}^{n} \gamma_j \hat{E}_K(c_j \lambda) \tag{10}$$

where γ_j is the weight of the expectation value with noise rate scaled as $c_j\lambda$. To do n steps at once, we have to follow 2 constraints:

1. $E^\star$ has to end up normalized at the end. So

$$\sum_{j=0}^{n} \gamma_j = 1$$

2. We need to cancel out n terms from RHS of equation 9 to improve our aprroximation. So

$$\sum_{j=0}^{n} \gamma_j c_j^k = 1 \text{ for } k = 1 \ldots n$$