Lanczos algorithm

Friday, March 11, 2022 1:50 PM

Why Lanczos algorithm?

- Local Hamiltonians describing qubit systems are SPARSE :

for nearest-neighbour interaction only :

$$\hat{H} = \sum_i \omega_i \hat{\sigma}_z + \sum_{\substack{i,j \\ i \neq j}} g \left(\hat{\sigma}_i^+ \hat{\sigma}_j^- + \hat{\sigma}_i^- \hat{\sigma}_j^+ \right)$$

$$\hat{H} = \begin{pmatrix} \omega_1 & g_{12} & & \emptyset \\ g_{12} & \omega_2 & \ddots & \\ & \ddots & \ddots & g_{N-1} \\ \emptyset & & g_{N-1} & \omega_N \end{pmatrix}$$

$\dim(H) = N \times N = 2^n \times 2^n$ with n qubits

- Full diagonalization does not make use of sparsity

↳ Calculating the ground state by diagonalizing H doesn't use sparsity

↳ Calculating time evolution requires exponentiating H

$|\psi(t)\rangle = e^{-iHt} |\psi(0)\rangle$ if $\Delta t = t - 0$ is small enough

which can be done through diagonalization of H

$H = S^\dagger D S$ where D is a diagonal matrix
S is unitary

then

$$e^{-iHt} = \mathbb{1} - iHt - \frac{1}{2} H^2 t^2 + \dots$$

$$= \mathbb{1} - i (S^\dagger D S) t - \frac{1}{2} (S^\dagger D S)(S^\dagger D S) t^2 + \dots$$

$$= \mathbb{1} - i (S^\dagger D S) t - \frac{1}{2} S^\dagger D^2 S \, t^2 + \dots$$

$$= S^{+} \left(1 - i\, D t - \frac{1}{2} D^2 t^2 + \dots \right) S$$

$$= S^{+} e^{-iDt} S$$

We can calculate the matrix exponential by first diagonalizing H and then exponentiating a diagonal matrix D (which means exponentiating each of the diagonal elements)

Time evolution doesn't use the sparsity of H

* Matrix - vector multiplication can use sparsity

for example, create a matrix with same # of rows of original matrix and a tuple for each non-zero matrix element which stores the column index and the value.

For each row of the matrix (each can be done in parallel) sum the product of the non-zero values in the marked column by the corresponding element in the vector's rows

*

$$A = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 4 \end{bmatrix} \qquad x = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

$$A\, x = \begin{bmatrix} 1\cdot 1 + 0\cdot 2 + 0\cdot 3 \\ 0\cdot 1 + 2\cdot 2 + 0\cdot 3 \\ 0\cdot 1 + 0\cdot 2 + 4\cdot 3 \end{bmatrix} \rightarrow 15 \text{ operations}$$

$$A_{sparse} = \begin{bmatrix} (0,1) \\ (1,2) \\ (2,4) \end{bmatrix}$$

$$A_{sparse}\, x = \begin{bmatrix} 1\cdot 1 \\ 2\cdot 2 \\ 4\cdot 3 \end{bmatrix} \rightarrow 3 \text{ operations}$$

↳ Time-evolution can benefit from sparse matrices

$$|\psi(t)\rangle = e^{-iHt}|\psi_0\rangle$$

$$= \sum_n \frac{(-it)^n}{n!} H^n |\psi_0\rangle$$

While H^n becomes less and less sparse as $n \to \infty$, we can instead split the sum above as

$|u_0\rangle = |\psi_0\rangle$, $|u_1\rangle = H|u_0\rangle$, $|u_2\rangle = H|u_1\rangle$, ...

so that we always multiply a sparse matrix with a vector

$$= \sum_n \frac{(-it)^n}{n!} |u_n\rangle$$

⚠ it only works for $t \to 0$, errors accumulate fast!

Krylov subspace

Given a matrix H and a vector $|\psi_0\rangle$, the k-order Krylov subspace is defined as the space spanned by

$$K_r = \{ |\psi_0\rangle, H|\psi_0\rangle, H^2|\psi_0\rangle, H^3|\psi_0\rangle, \dots, H^{k-1}|\psi_0\rangle \}$$

$$K_r \{ |u_0\rangle, |u_1\rangle, \dots, |u_{k-1}\rangle \}$$

Using Gram-Schmidt we can find an orthonormal basis for the space:

$$|\tilde{\phi}_0\rangle = |u_0\rangle \to |\phi_0\rangle = \frac{|\tilde{\phi}_0\rangle}{\langle\tilde{\phi}_0|\tilde{\phi}_0\rangle}$$

$$|\tilde{\phi}_1\rangle = |u_1\rangle - (\langle\phi_0|u_1\rangle)|\phi_0\rangle \to |\phi_1\rangle = \frac{|\tilde{\phi}_1\rangle}{\langle\tilde{\phi}_1|\tilde{\phi}_1\rangle}$$

$\vdots$

$$|\tilde{\phi}_{k-1}\rangle = |u_{k-1}\rangle - \sum_{i=0}^{k-2} \left(\langle \phi_i | u_{k-1}\rangle\right)|\phi_i\rangle \rightarrow |\phi_{k-1}\rangle = \frac{|\tilde{\phi}_{k-1}\rangle}{\langle \tilde{\phi}_{k-1}|\tilde{\phi}_{k-1}\rangle}$$

Using the basis $|\phi\rangle$, we can define the matrix Q

$$Q = \begin{bmatrix} \phi_0 & \phi_1 & \phi_2 & \cdots & \phi_{k-1} \\ \vdots & \vdots & \vdots & & \vdots \end{bmatrix} \qquad \dim(Q) = N \times k$$

$$\alpha_0 = \langle \phi_0 | \underbrace{H|\phi_0\rangle}_{|\phi_1\rangle}$$

Then $\quad H \simeq Q T Q^{+}$

$$\beta_0 = \langle \phi_1 | H|\phi_0\rangle = \langle \phi_1 | \phi_1\rangle$$

where $T = \begin{pmatrix} \alpha_0 & \beta_0 & & \\ \beta_0 & \alpha_1 & \beta_1 & \\ & \beta_1 & \ddots & \beta_{k-2} \\ & & \beta_{k-2} & \alpha_{k-1} \end{pmatrix}$ is a real, symmetric tridiagonal matrix

$$\dim(T) = k \times k$$

↳ For - Time evolution

$$|\psi\rangle = e^{-iHt}|\psi_0\rangle$$

$$= e^{-iQTQ^{+}t}|\psi_0\rangle$$

$$= Q\, e^{-iTt}\, Q^{+}|\psi_0\rangle$$

$$= Q\, V\, e^{-i\,\mathrm{diag}(T)}\, V^{+}\, Q^{+}|\psi_0\rangle$$

we only need to diagonalize T which has dimension $k \ll N$

$\uparrow$ size of the krylov subspace

Overview of the algorithm

set $\quad |\phi_0\rangle = \dfrac{|\psi_0\rangle}{\langle \psi_0|\psi_0\rangle} \quad , \quad \beta_{-1} = 0$

for $l = 0$ to $k-1$:

$|v\rangle = H|\phi_j\rangle \rightarrow HQ$

$\alpha_j = \langle\phi_j|v\rangle \rightarrow$ comes from $Q^+(HQ) = T$ on the diagonal

$|v_1\rangle = |v\rangle - \alpha_j|\phi_j\rangle - \beta_{j-1}|\phi_{j-1}\rangle \rightarrow$ Gram - Schmidt

$\beta_j = \langle v_1|v_1\rangle \rightarrow$ from $Q^+(HQ) = T$ off diagonal

if $\beta_j = 0 \rightarrow$ STOP: found a smaller subspace than k

else:

$$|\phi_{j+1}\rangle = \frac{|v_1\rangle}{\beta_j}$$

store $\alpha_j, \beta_j, |\phi_j\rangle$

$$\rightarrow Q = \left[|\phi_0\rangle \;\vdots\; |\phi_1\rangle \;\vdots\; \cdots |\phi_{k-1}\rangle \right]$$

$$T = \begin{bmatrix} \alpha_0 & \beta_0 & & \emptyset \\ \beta_0 & \alpha_1 & & \\ & & \alpha_2 & & \beta_{k-2} \\ \emptyset & & & \beta_{k-2} & \alpha_{k-1} \end{bmatrix}$$