Reflections in Hilbert Space II: Szegedy's scheme for Markov chain quantization

James Daniel Whitfield[*][†]

March 20, 2012

> The cops have plans, Gordon's got plans. You know, they're schemers. Schemers trying to control their little worlds. I'm not a schemer.
> –*Joker, in The Dark Knight*

The geometric picture developed in previous lecture for Grover's search can be applied to any pair of projectors. This is the key to most of the quantum speed-ups that are quadratic. Some of these generalization and applications of arbitrary reflection operators were pointed out in [1]. Szegedy published a pair of papers[2, 3] that contained many of the same results, however builds upon previous work by pointing out a correspondence between projection on to subspaces associated with Markov chains. This is the focus of this lecture.

Szegedy's scheme is an important generalization of Grover's algorithm because it simplifies access to the promised quadratic speed up of quantum computation and, like Grover's search, employs the product of two reflections for unitary dynamics. The most important technical contribution of Szegedy's scheme is the spectral decomposition which is covered in the next lecture. The final portion of this lecture is a review of the singular value decomposition used heavily in the spectral decomposition.

1 Markov Chain Quantization

1.1 Markov Chains

Consider a series of random variables indexed by a integer parameter (usually associated with time), $\{X_i\}$. The probability distributions of these random variables is written like, $P(X_{n-1}, X_{n-2}, \ldots, X_3, X_2, X_1)$. The conditional probability for X_n given X_{n-1} through X_1 is $P(X_n|X_{n-1}, X_{n-2}, \ldots, X_1)$. If it is the case that $P(X_n|X_{n-1}, X_{n-2}, \ldots, X_1) = P(X_n|X_{n-1})$, the chain of random variable is considered Markovian. In words, the probability of each random

[*]Columbia University, NEC Labs America, Institute for Scientific Interchange

[†]*Current address:* Institute for Scientific Interchange, Via Alassio 11/c, 10126 Torino - Italy

variable depends only on the previous random variable. Then using Baye's rule, $P(X_n|X_{n-1}) = \frac{P(X_n,X_{n-1}}{P(X_{n-1})}$, we can write

$$P(X_n = s_i) = \sum_k P(X_n = s_i|X_{n-1} = s_k)P(X_{n-1} = s_k) \tag{1}$$

where s_i and s_k are realizations from the sample space of the random variables.

The Markov transition matrix for chain M, is $M_{ij} = P(X_n = s_j|X_{n-1} = s_i) = prob(s_i \to s_j)$. With this indexing convention, matrix multiplication of probability vectors is done from the left so that probability vector u, with $|u|_1 = \sum_i u_i = 1$, transforms as $uM = u'$. Note clearly that the indexing convention is not the same in all texts and had we chosen $\tilde{M}_{ij} = P(X_n = s_i|X_{n-1} = s_j)$ matrix multiplication would occur from the right. The Markov chain is stationary if $P(X_n|X_{n-1}) = P(X_m|X_{m-1})$ for all n, m. We will only consider stationary Markov processes.

From the definitions given above, $\sum_j M_{ij} = 1$, and $Mc = c$ with $c_j = 1$ for all j. Furthermore, the Perron-Frobenius theorem states that there is an (left) eigenvector with eigenvalue one with all positive entries, and it is the only such eigenvector if the chain is irreducible or ergodic (i.e. the underlying graph is connected).

1.2 Szegedy's quantization scheme

Szegedy's Markov chain quantization scheme can be found in [2, 3]. It requires two Markov chains to construct a reflection-type operator as in the Grover search algorithm. Each reflection is associated with one of the two Markov chains. We will call them P and Q. Suppose the chains act over some state space $\Omega_A = \{s_i\}$ and $\Omega_B = \{s_i\}$ where it is not required that the state spaces have the same dimension.

To begin we define the transition vectors as

$$|\phi_x\rangle = \sum_y \sqrt{p_{xy}}|x\rangle|y\rangle \tag{2}$$

$$|\psi_y\rangle = \sum_x \sqrt{q_{yx}}|x\rangle|y\rangle \tag{3}$$

These transition vectors are defined on a vector space $H_A \otimes H_B$ where the basis $\{|x\rangle\}$ ($\{|y\rangle\}$) of H_A (H_B) corresponds to states of Ω_A (Ω_B). Since each row of a Markov chain sums to unity and each transition vector corresponds to a row of the Markov chain, the transition vectors *correspond* to probability distributions. For now we accept the definitions and more motivation will be provided as we proceed.

Using the fact that the row sum to unity,

$$\begin{aligned}
\langle\phi_{x'}|\phi_x\rangle &= \left(\sum_y \sqrt{p_{xy}}\langle x|\langle y|\right)\left(\sum_z \sqrt{p_{x'z}}|x'\rangle|z\rangle\right) \\
&= \sum_{yz} \delta_{yz}\sqrt{p_{xy}p_{x'z}}\delta_{xx'} \\
&= \sum_y \sqrt{p_{xy}p_{x'y}}\delta_{xx'} = (\sum_y p_{xy})\delta_{xx'} \\
&= \delta_{xx'}
\end{aligned}$$

Similarly, one can show that $\{\psi_y\}$ also form an orthonormal set and that the overlap between the two spaces is

$$\langle\phi_x|\psi_y\rangle = \sqrt{p_{xy}q_{yx}}. \tag{4}$$

Now we are in a position to define the Szegedy evolution operator W. First, consider two half projectors, $B = \sum_y |\psi_y\rangle\langle y|$ and $A = \sum_x |\phi_x\rangle\langle x|$. In the next lecture these will be useful for obtaining the spectral decomposition of W. For now, they define two projectors via $AA^\dagger$ and $BB^\dagger$. Performing a product of reflection just as in Grover's search provides the evolution operator.

$$W = (2AA^\dagger - \mathbf{1})(2BB^\dagger - \mathbf{1}) \tag{5}$$

2 The Hilbert space

Let us compare the Szegedy operator with the Grover operator, $-R_sR_w$, from the first lecture. With the Grover operator, we made progress be considering the space spanned by the w and s (the states about which the reflections R_s and R_w are made). The space orthogonal to these states was trivial and we were able to ignore it.

Similarly, the Szegedy operator partitions the Hilbert space into different region based upon its action on the states. The Hilbert space is a tensor of vector spaces corresponding to the two sample spaces. Since the transition vectors live on the combined space we can consider the Hilbert space as the edge space of the Markov chains with $\mathcal{E} = \{|v_1\rangle|v_2\rangle : |v_1\rangle \in H_A, |v_2\rangle \in H_B\}$. Visually, this is depicted in fig. 1.

The trivial space is spanned by states that are orthogonal to both $\{|\psi_y\rangle\}$ and $\{|\phi_x\rangle\}$. The operator W has no action on these states. In the inactive spaces, each application of W changes the sign of the states it acts on. More specifically, if $|a\rangle$ is in H_A but not in H_B then

$$W|a\rangle = (2AA - 1)(2BB - 1)|a\rangle = (2AA - 1)(-|a\rangle = -|a\rangle$$

The active space spanned by the states, $\{|\psi_y\rangle\}$ and $\{|\phi_x\rangle\}$ is where non-trivial evolution under W occurs.

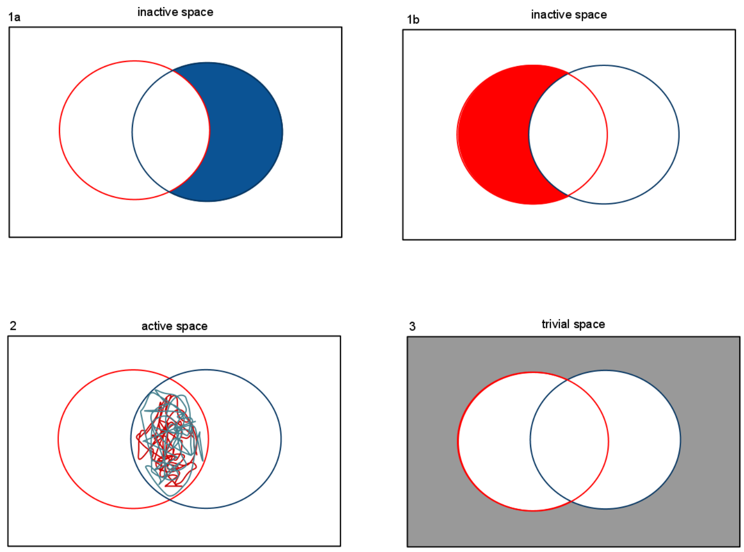

Figure 1: Visual representation of the Hilbert space that the Szegedy operator acts upon. In 1a and 1b, the inactive spaces where each use of W only changes the sign of the state is shown. In 3, the trivial space, where W has no action, is shown. Finally, in 2, the active space is shown. This is where the interesting dynamics occur.

Just as in the analysis of Grover's operator, the trivial space can be characterized by, first, orthogonalizing the space spanned by both sets of transition vectors. Next, given an arbitrary vector we can subtract the projection onto the transition vector space to obtain a member of the trivial space. By repeating this process and orthogonalizing the obtained set we have a basis for the trivial space.

The inactive space can be characterized by selecting any vector, subtracting the projection onto the trivial space and the projection onto either H_A or H_B. Since the active space has nontrivial trajectories, we will diagonalize it in the next lecture following a calculation reproduced from [3]. We close this lecture with a review of the singular value decomposition and the introduction of canonical angles which will prove useful in the next lecture.

3 Singular value decomposition

The singular value decomposition is similar to the eigenvalue decomposition; however, it can be applied to matrices that are not square. Singular values are frequently used to characterize norms of matrices. For instance, the operator norm is the largest singular value and the trace norm is the sum of the singular values.

The singular value decomposition of T says that

$$T = V\Sigma W^\dagger \tag{6}$$

with the following spatial decomposition: $[n\times m] = [n\times n][n\times m][m\times m]$. Here $[n\times m]$ denotes the space of matrices with n columns and m rows. V and W are unitary and Σ is diagonal. The diagonal elements of Σ are the singular values and, they are uniquely specified by T.

The singular values of matrix T are the eigenvalues of $\sqrt{T^\dagger T}$. Since $T^\dagger T$ is positive semi-definite, the square root function is well defined. Additionally, since $T^\dagger T$ is normal ($[T, T^\dagger] = 0$) we have $T^\dagger T = V\Lambda V^\dagger$ with Λ a diagonal matrix and V a unitary matrix. We define the singular matrix as $\Sigma \equiv \sqrt{\Lambda}$.

For arbitrary matrix, T, we now give constructions for the input and output unitary matrices, V and W, which will prove that the SVD exists for all matrices. If T is non-singular, we can invert the diagonal singular value matrix, Σ^{-1} by taking the reciprocal of the singular values. Letting $W = T^\dagger V\Sigma^{-1}$, W is unitary and $T = V\Sigma W^\dagger$.

When T is singular, the inverse of Σ is no longer available as some of the singular values are zero. Instead, the pseudo-inverse is taken where only the inverse of non-zero diagonal elements is used. As before, the output matrix W is defined using the pseudo-inverse. To satisfy the unitary constraint, arbitrary orthonormal vectors are included as columns of W. Thus, the singular value decomposition exists for all matrices regardless of shape and the matrices need not satisfy special properties.

The vector form of the SVD can be easily seen as $TW = V\Sigma$ and $V^\dagger T = \Sigma W^\dagger$ imply

$$\begin{aligned} T|w_i\rangle &= \sigma_i|v_i\rangle \quad (7)\\ \langle v_i|T &= \langle w_i|\sigma_i \quad (8)\end{aligned}$$

The vectors $\{|v_i\rangle\}$ are called the left singular vectors and $\{|w_i\rangle\}$, the right singular vectors.

3.1 Normal matrices

If matrix A is normal ($A^\dagger A = AA^\dagger$), then the spectral decomposition exist for A that is $A = SDS^\dagger$ for some S unitary and D diagonal. The singular value decomposition is

$$
\begin{aligned}
A &= SDS^\dagger = S \begin{bmatrix} |\lambda_1|e^{i\varphi_1} & & & \\ & |\lambda_2|e^{i\varphi_2} & & \\ & & \ddots & \\ & & & |\lambda_n|e^{i\varphi_n} \end{bmatrix} S^\dagger && (9) \\
&= S \begin{bmatrix} |\lambda_1| & & & \\ & |\lambda_2| & & \\ & & \ddots & \\ & & & |\lambda_n| \end{bmatrix} \begin{bmatrix} e^{i\varphi_1} & & & \\ & e^{i\varphi_2} & & \\ & & \ddots & \\ & & & e^{i\varphi_n} \end{bmatrix} S^\dagger && (10) \\
&= S\Sigma D S^\dagger && (11) \\
&= S\Sigma (SD^\dagger)^\dagger = V\Sigma W^\dagger && (12)
\end{aligned}
$$

3.2 Canonical angles

Canonical angles determine how far apart two spaces are from one another. It is a useful measure since it is unitarily invariant. This implies that the bases chosen to describe the two spaces do not matter. We will only consider spaces of equal dimension but it is not difficult to generalize to non-equal dimensions.

From the bases $\{|i\rangle\}$ and $\{|j\rangle\}$ for spaces $\mathcal{H}_A$ and $\mathcal{H}_B$, we define the unitary matrices A and B such that the columns of A (B) are the vectors $\{|i\rangle\}(\{|j\rangle\})$. The singular values of $C = A^\dagger B$ can be bounded by unity due to the submultiplicative of matrix norms ($||AB|| \leq ||A||\ ||B||$). Since A and B are unitary, their singular values (the absolute value of the eigenvalues) are 1. Therefore, $||C|| \leq ||A^\dagger||\ ||B|| = 1 \cdot 1 = 1$ where we have used the operator norm (largest singular value).

Thus, all the singular values of C are between zero and one, so for each singular value σ_i we can define an angle as $\cos\theta_i = \sigma_i$. This angle θ_i is called the canonical angle. The canonical angles will show up in the next lecture when describing the action of the of the Szegedy operator.

References

[1] G. Brassard, P. Hoyer, M. Mosca, and A. Tapp. Quantum amplitude amplification and estimation. *arxiv:quant-ph/0005055*, 2000.

[2] M. Szegedy. Spectra of quantized walks and a $\sqrt{}$ rule. *arXiv*, 2004. http://arxiv.org/pdf/quant-ph/0401053v1.pdf.

[3] M. Szegedy. Quantum speed-up of Markov chain based algorithms. In *Proc. 45th Ann. IEEE Symp. Found. Comput. Sci.*, pages 32–41, 2004. http://www.cs.rutgers.edu/ szegedy/PUBLICATIONS/walk_focs.pdf.