EECS 16B Designing Information Devices and Systems II
Fall 2021 Note 18: Justifying PCA Via The Eckart-Young Theorem

In this note we will discuss the proof of the so-called *Eckart-Young theorem*, which is a result we put off in the last note for the sake of brevity, since the proof is rather lengthy.

As a reminder, the Eckart-Young theorem states that the best rank-k approximation to a matrix A is the approximation $\widehat{A}$ generated by truncating all but the top k singular values. To be precise, the theorem states that if

$$A = U\Sigma V^\top = \sum_{i=1}^{r} \sigma_i \vec{u}_i \vec{v}_i^\top \tag{1}$$

is the SVD of A, then

$$A_k = \sum_{i=1}^{k} \sigma_i \vec{u}_i \vec{v}_i^\top \tag{2}$$

satisfies

$$A_k = \underset{\substack{\widehat{A}\in\mathbb{R}^{m\times n} \\ \text{rank}(\widehat{A})\leq k}}{\text{argmin}} \left\|A - \widehat{A}\right\|_F^2. \tag{3}$$

1 Frobenius Norm

How can we prove this result? First, we should take the time to establish some properties of the Frobenius norm.

Consider an $m \times n$ matrix M. As stated above, the Frobenius norm of this matrix is defined to be

$$\|M\|_F^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} M_{ij}^2. \tag{4}$$

First, let's look at how the columns of M relate to this quantity. Let the columns of M be m-dimensional vectors $\vec{v}_i$ such that

$$M = \begin{bmatrix} | & & | \\ \vec{v}_1 & \cdots & \vec{v}_n \\ | & & | \end{bmatrix}. \tag{5}$$

Observe that the squared norm of a particular vector is

$$\left\|\vec{v}_j\right\|^2 = \sum_{i=1}^{m} M_{ij}^2. \tag{6}$$

In other words, it is the sum of squares of all the elements in the j^{th} column. Summing this quantity over all the columns, we obtain the sum of squares of all the elements in M, which is the squared Frobenius norm.

Expressed algebraically, it is the case that

$$\|M\|_F^2 = \sum_{i=1}^{m}\sum_{j=1}^{n} M_{ij}^2 = \sum_{j=1}^{n}\left(\sum_{i=1}^{m} M_{ij}^2\right) = \sum_{j=1}^{n}\left\|\vec{v}_j\right\|^2. \tag{7}$$

A similar result can be derived in an exactly analogous manner for the rows of M.

Now, we will look at how the Frobenius norm of a matrix varies under an orthonormal change of basis. Let Q be a matrix with m orthonormal columns, such that $Q^\top Q = I_{m\times m}$. Observe that, using the above result,

$$\begin{aligned}
\|QM\|_F^2 &= \left\| \begin{bmatrix} | & & | \\ Q\vec{v}_1 & \cdots & Q\vec{v}_n \\ | & & | \end{bmatrix} \right\|_F^2 && (8)\\
&= \sum_{j=1}^{n}\left\|Q\vec{v}_j\right\|^2 && (9)\\
&= \sum_{j=1}^{n}\vec{v}_j^\top Q^\top Q\vec{v}_j && (10)\\
&= \sum_{j=1}^{n}\vec{v}_j^\top \vec{v}_j && (11)\\
&= \sum_{j=1}^{n}\left\|\vec{v}_j\right\|^2 && (12)\\
&= \|M\|_F^2, && (13)
\end{aligned}$$

so the Frobenius norm of a matrix remains unchanged under pre-multiplication by Q. One way to interpret the above algebraic result is to notice that pre-multiplying by Q doesn't change the norm of any of the columns of M, so its Frobenius norm should similarly remain unchanged. It is interesting to note that here, we don't actually need Q to be square — that never came up in the above derivation. We just need orthonormal columns.

In a very similar manner, it can be shown that post-multiplication by a matrix $Q^\top$ with orthonormal rows again does not change the Frobenius norm. A way of seeing this is to simply take transposes everywhere in the above calculation, recognizing that the Frobenius norm of a matrix does not change if we take the transpose.

Now, combining the above observations with our knowledge of the SVD, we find that, given a matrix A with SVD $A = U\Sigma V^\top$,

$$\begin{aligned}
\|A\|_F &= \left\|U\Sigma V^\top\right\|_F && (14)\\
&= \left\|U^\top U\Sigma V^\top V\right\|_F && (15)\\
&= \|\Sigma\|_F, && (16)
\end{aligned}$$

since U and V are both square[1] orthonormal matrices, so pre-multiplying by $U^\top$ and post-multiplying by

[1] Here, we are using the full form of the SVD, not the compact form.

V both do not change the Frobenius norm, as in both cases their rows and columns are all orthonormal. Writing out the Frobenius norm of Σ in terms of the singular values explicitly, we find that

$$\|A\|_F = \sqrt{\sum_{i=1}^{r} \sigma_i^2}. \tag{17}$$

This seems like a useful result that we should hang on to.

2 Reducing to an easier case via the SVD

Now, let's get back to the question of low-rank approximations. As stated before, we wish to choose a rank at most k matrix $\widehat{A}$ that minimizes $\left\|A - \widehat{A}\right\|_F$. Looking at the SVD of A and applying simplifications very similar to those derived in the previous section, pre-multiplying by $U^\top$ and post-multiplying by V, we see that

$$\begin{aligned}
\left\|A - \widehat{A}\right\|_F &= \left\|U\Sigma V^\top - \widehat{A}\right\|_F && (18)\\
&= \left\|U^\top U\Sigma V^\top V - U^\top \widehat{A} V\right\|_F && (19)\\
&= \left\|\Sigma - U^\top \widehat{A} V\right\|_F. && (20)
\end{aligned}$$

Let $X = U^\top \widehat{A} V$, for convenience. Observe that, since U and V are both invertible, given any X we can choose a corresponding $\widehat{A}$, and vice-versa. Furthermore, the rank of $\widehat{A}$ and X are the same, so X has rank no more than k. Thus, the problem reduces to choosing an X of rank no more than k that minimizes

$$\|\Sigma - X\|_F. \tag{21}$$

Essentially, the SVD tells us that all we really need to understand is the diagonal case. How can we best approximate a diagonal matrix with sorted entries down the diagonal using a rank at most k matrix?

Let's write out Σ and X explicitly, to try and get a better idea of what remains for us to prove. We wish to choose columns $\vec{x}_i$ that minimize

$$\begin{aligned}
&\left\| \begin{bmatrix} \Sigma_r & 0_{r\times(n-r)} \\ 0_{(m-r)\times r} & 0_{(m-r)\times(n-r)} \end{bmatrix} - X \right\|_F^2 && (22)\\
=&\left\| \begin{bmatrix} \sigma_1\vec{e}_1 & \cdots & \sigma_r\vec{e}_r & \vec{0} & \cdots & \vec{0} \end{bmatrix} - \begin{bmatrix} | & & | & | & & | \\ \vec{x}_1 & \cdots & \vec{x}_r & \vec{x}_{r+1} & \cdots & \vec{x}_n \\ | & & | & | & & | \end{bmatrix} \right\| && (23)
\end{aligned}$$

We will use the notation $\vec{e}_i$ here to represent the i^{th} column of the identity matrix $I_{m\times m}$ throughout the note.

Expressing the squared Frobenius norm as the sum of the squared norms of the columns, we see that we are tasked with picking columns $\vec{x}_i$ minimize

$$\|\sigma_1\vec{e}_1 - \vec{x}_1\|^2 + \cdots + \|\sigma_r\vec{e}_r - \vec{x}_r\|^2 + \left\|\vec{0} - \vec{x}_{r+1}\right\|^2 + \cdots + \left\|\vec{0} - \vec{x}_n\right\|^2 \tag{24}$$

$$= \sum_{i=1}^{r} \|\sigma_i \vec{e}_i - \vec{x}_i\|^2 + \sum_{i=r+1}^{n} \|\vec{x}_i\|^2, \tag{25}$$

while keeping the rank of X less than or equal to some given constant k. First, we should pick

$$\vec{x}_{r+1} = \vec{x}_{r+2} = \cdots = \vec{x}_n = \vec{0}, \tag{26}$$

since making them nonzero would hurt us by increasing the sum of norms that we are trying to minimize.

But what about the $\vec{x}_1, \ldots, \vec{x}_r$? There are two cases here.

If $k \geq r$, then we can pick $\vec{x}_i = \sigma_i \vec{e}_i$ for $i = 1, \ldots, r$. Then X is a perfect rank $\leq k$ approximation for Σ, i.e., $X = \Sigma$ and $\operatorname{rank} X \leq k$. This can be turned into a perfect approximation for A, i.e., $\widehat{A} = A$. This confirms our claim that the best rank $\leq k$ approximation for A is just A, in the case that $k \geq r$. We are now done for the case $k \geq r$. So from now on let us assume $k < r$.

The other case, $k < r$, is more interesting. Because of the constraint on the rank of X, we can't make all $\vec{x}_i$ equal to the corresponding columns of Σ, as then the rank of X would be r, and in particular greater than k.

One intuitive guess could be to remove the k largest singular values since they are affecting the norm the most. Then we would pick $\vec{x}_1 = \sigma_1 \vec{e}_1, \vec{x}_2 = \sigma_2 \vec{e}_2, \ldots, \vec{x}_k = \sigma_k \vec{e}_k$ and set the rest of the $\vec{x}_i = \vec{0}$ for $i > k$. This definitely makes X have rank k, and turns out to be the correct answer. But how can we be sure that there isn't a better choice out there? This requires a proof, which we will do in the next section.

3 The Projection Perspective

In the previous section, we used the SVD to distill our problem down to the diagonal case. The problem is: how can we choose an X of rank no more than k that minimizes

$$\|\Sigma - X\|_F? \tag{27}$$

We saw that what matters here is choosing the r vectors $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_r$ that are matched up with the non-zero parts of Σ in the above subtraction, since the rest of X should be all zeros. We have a guess as to what the optimal choice should be, but we need to prove that it is indeed optimal.

The projection perspective approaches this problem by stepping back a bit. Dealing with the constraint on the rank of X is annoying. What does this condition really mean from a geometric point of view? It means that all these r vectors $\vec{x}_1, \vec{x}_2, \ldots, \vec{x}_r$ must lie in a k-dimensional subspace. This is what rank means.

Let's reason about the problem in terms of a basis for this subspace. Imagine that we are already given this subspace as the span of some $m \times k$ matrix Q. What kind of matrix should we ask for? We know that orthonormal matrices are comparatively easy to work with, and any other matrix with linearly independent columns can be orthonormalized, so for convenience, let's assume that all the columns of Q have been orthonormalized, so $Q^\top Q = I_{k \times k}$. Then what should each of the $\vec{x}_i$ be, assuming they are constrained to lie in the column space of Q?

Well, once we're given Q, we simply need to choose each $\vec{x}_i$ to minimize the term it is involved in — specifically, each $\vec{x}_i$ should minimize

$$\|\sigma_i \vec{e}_i - \vec{x}_i\|^2. \tag{28}$$

And how can we choose such an $\vec{x}_i$? We know this from 16A! It is simply the projection of $\sigma_i \vec{e}_i$ onto the

column space of Q, which from least squares is just

$$\vec{x}_i = Q(Q^\top Q)^{-1}Q^\top(\sigma_i \vec{e}_i) = QQ^\top(\sigma_i \vec{e}_i) = \sigma_i QQ^\top \vec{e}_i, \tag{29}$$

simplifying by recalling that $Q^\top Q = I_{k\times k}$, as we chose the columns of Q to be an orthonormal basis.

This means that finding the right orthonormal Q is enough to solve our problem. In hindsight, this is not surprising since our initial motivation was to discover a subspace that best captures the data.

4 Finding the Best Q

So our problem has been simplified further. Rather than trying to find an arbitrary $m \times n$ matrix X with rank $\leq k$, we can instead find an $m \times k$ orthonormal matrix Q, and then compute $\vec{x}_i$ by projecting each of the columns of Σ onto the column space of Q.

Since we're now working with Q, rather than X, it makes sense to write the quantity we're interested in minimizing in terms of Q. Substituting in our least-squares solution for each of the $\vec{x}_i$ into our earlier expression for the squared Frobenius norm, and setting all the $\vec{x}_i = \vec{0}$ for all $i > r$ (as we argued would be optimal), we find that the new quantity to minimize becomes

$$\sum_{i=1}^{r} \|\sigma_i \vec{e}_i - \vec{x}_i\|^2 + \sum_{i=r+1}^{n} \|\vec{x}_i\|^2 \tag{30}$$

$$= \sum_{i=1}^{r} \left\|\sigma_i \vec{e}_i - \sigma_i QQ^\top \vec{e}_i\right\|^2 + \sum_{i=r+1}^{n} \left\|\vec{0}\right\|^2 \tag{31}$$

$$= \sum_{i=1}^{r} \sigma_i^2 \left\|\vec{e}_i - QQ^\top \vec{e}_i\right\|^2. \tag{32}$$

Can we simplify this expression somewhat? Notice that $\left(\vec{e}_i - QQ^\top \vec{e}_i\right) \perp QQ^\top \vec{e}_i$, since we know from 16A that the least squares projection is orthogonal to the residual error vector. Thus, by the Pythagorean theorem, we can write

$$\|\vec{e}_i\|^2 = \left\|\vec{e}_i - QQ^\top \vec{e}_i\right\|^2 + \left\|QQ^\top \vec{e}_i\right\|^2 \implies \left\|\vec{e}_i - QQ^\top \vec{e}_i\right\|^2 = \|\vec{e}_i\|^2 - \left\|QQ^\top \vec{e}_i\right\|^2. \tag{33}$$

Making this substitution in our earlier expression, noticing that $\|\vec{e}_i\| = 1$ and rearranging, our problem reduces to finding an orthonormal Q that minimizes the quantity

$$\sum_{i=1}^{r} \sigma_i^2 \left(\|\vec{e}_i\|^2 - \left\|QQ^\top \vec{e}_i\right\|^2\right) = \sum_{i=1}^{r} \sigma_i^2 - \sum_{i=1}^{r} \sigma_i^2 \left\|QQ^\top \vec{e}_i\right\|^2 \tag{34}$$

$$= \sum_{i=1}^{r} \sigma_i^2 - \sum_{i=1}^{r} \left\|QQ^\top \sigma_i \vec{e}_i\right\|^2 \tag{35}$$

$$= \|\Sigma\|_F^2 - \left\|QQ^\top \Sigma\right\|_F^2 \tag{36}$$

Notice that we have broken the quantity we are minimizing into two summations. The first summation is simply the squared Frobenius norm of Σ, which does not depend on Q. The second is the negation of the

squared Frobenius norm of $QQ^\top\Sigma$. Thus, in order to minimize the overall quantity, we should aim to choose our Q to *maximize* $\left\|QQ^\top\Sigma\right\|_F^2$, since we are subtracting this quantity from another fixed value.

5 Working with this even simpler problem

Cool! We've managed to simplify our problem further, by taking advantage of properties of projections and orthonormal bases. Recall from earlier that $\|QM\|_F^2 = \|M\|_F^2$ when Q has orthonormal columns. Then we see that the quantity we are trying to maximize can be slightly simplified as

$$\left\|QQ^\top\Sigma\right\|_F^2 = \left\|Q^\top\Sigma\right\|_F^2. \tag{37}$$

Be aware that, since Q is not square, $QQ^\top \neq I_{m\times m}$, as the rows are not necessarily mutually orthogonal. So we can't use this property to make any further simplifications.

How can we show that our guessed solution (choosing the first k columns of the identity for Q) is indeed the best choice for maximizing (37)? That choice would give $\left\|Q^\top\Sigma\right\|_F^2 = \sum_{j=1}^k \sigma_j^2$. One way to show that this is optimal is to show that we cannot possibly do any better than that.

To do this, let's just expand this product out directly. For notational convenience, let the columns of Q be $\vec{q}_i$, and so

$$Q^\top = \begin{bmatrix} - & \vec{q}_1^\top & - \\ - & \vec{q}_2^\top & - \\ & \vdots & \\ - & \vec{q}_k^\top & - \end{bmatrix}. \tag{38}$$

Remember that each of the $\vec{q}_i$ are m-dimensional orthonormal vectors.

Now, writing out our product, we obtain

$$\left\|Q^\top\Sigma\right\|_F^2 = \left\| \begin{bmatrix} - & \vec{q}_1^\top & - \\ - & \vec{q}_2^\top & - \\ & \vdots & \\ - & \vec{q}_k^\top & - \end{bmatrix} \begin{bmatrix} \sigma_1\vec{e}_1 & \sigma_2\vec{e}_2 & \cdots & \sigma_r\vec{e}_r & \vec{0} & \cdots & \vec{0} \end{bmatrix} \right\|_F^2 \tag{39}$$

$$= \sum_{i=1}^k \sum_{j=1}^r \left(\vec{q}_i^\top \vec{e}_j\right)^2 \sigma_j^2 \tag{40}$$

$$= \sum_{i=1}^k \sum_{j=1}^r Q_{ji}^2 \sigma_j^2 \tag{41}$$

$$= \sum_{j=1}^r \sigma_j^2 \sum_{i=1}^k Q_{ji}^2 \tag{42}$$

Thus, we have expressed our desired Frobenius norm as a linear combination of the squared singular values σ_j^2. For convenience and to focus our attention on what is going to matter, let the coefficients of our linear

combination be denoted as

$$d_j = \sum_{i=1}^{k} Q_{ji}^2, \tag{43}$$

where d_j is defined to be the squared norm of the jth row of Q. Then we can express the term we want to maximize as

$$\left\| Q^\top \Sigma \right\|_F^2 = \sum_{j=1}^{r} \sigma_j^2 d_j. \tag{44}$$

Now, observe that the Frobenius norm of $Q^\top$ itself is $\|Q\|_F = k$, since Q has k unit-norm columns. Expressing this quantity in terms of the d_j, we see that

$$\sum_{j=1}^{m} d_j = \left\| Q^\top \right\|_F^2 = \|Q\|_F^2 = k. \tag{45}$$

Furthermore, notice that, since they are defined as a sum of squares in (43), each of the $d_j \geq 0$. As $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_r$, we could try to upper-bound the quantity we are trying to maximize by choosing $d_1 = k$ and setting the other $d_j = 0$.

That would give $k\sigma_1^2$ which is bigger than our guessed optimal solution. But is this extreme choice of d_j actually achievable by some Q? The challenge is that we know a lot about the columns of Q since they are orthonormal. But d_j is related to the j^{th} *row* of Q. How can we understand anything about the rows of Q?

If Q were square, then the fact that $Q^\top Q = I$ would tell us that $Q^\top$ is the inverse of Q and so $Q^\top$ would also be orthonormal. But Q is decidedly not square, so what can we do?

Observe that the k columns of Q can be extended using Gram-Schmidt to form a full orthonormal basis for $\mathbb{R}^m$. Let this extended basis be

$$\widehat{Q} = \begin{bmatrix} | & & | & | & & | \\ \vec{q}_1 & \cdots & \vec{q}_k & \vec{q}_{k+1} & \cdots & \vec{q}_m \\ | & & | & | & & | \end{bmatrix}. \tag{46}$$

By the properties of orthonormal square matrices, we know that in addition to the columns, the *rows* of $\widehat{Q}$ also form an orthonormal basis, and so are all of unit norm. Thus, looking at the j^{th} row, we can write

$$\sum_{i=1}^{m} \widehat{Q}_{ji}^2 = 1 \tag{47}$$

for all j.

This tells us that the rows of $\widehat{Q}$ can't be too large, and thus, truncating the sum by removing the terms corresponding to the $(k+1)^{\text{th}}$ column onwards, we see that

$$d_j = \sum_{i=1}^{k} Q_{ji}^2 \leq \sum_{i=1}^{m} \widehat{Q}_{ji}^2 = 1. \tag{48}$$

So in addition to requiring that they sum to k, we have established that each individual $0 \leq d_j \leq 1$ for all j.

Now, we have found a simpler problem inside our original problem.

6 Working with the inner problem

At this point, we have turned even the simplified problem into something[2] only involving numbers. From (44), we know we want to maximize $\sum_{j=1}^{r} \sigma_j^2 d_j$ over the choice of d_j that we know must satisfy certain constraints. From (45), we know their overall sum $\sum_{j=1}^{m} d_j = k$. We also know from (48), that $0 \leq d_j \leq 1$ for all j.

A consequence of this is that our previous overly greedy assignment of values to d_j is not possible if $k > 1$, since we must have $d_1 \leq 1 < k$. Instead, if we wish to maximize our linear combination of the σ_j^2, we should set

$$d_1 = d_2 = \cdots = d_k = 1 \tag{49}$$

and

$$d_{k+1} = d_{k+2} = \cdots = d_m = 0, \tag{50}$$

in order to place as much weight as possible on the largest squared singular values without violating the constraints on the d_j. This gives us $\sum_{j=1}^{k} \sigma_j^2$ as the best we can do.

Why is this the case? We can think of each d_j as telling us how much money we want to spend to buy product j. The store only has up to 1 liter of each product in stock. Each item costs one dollar per liter. The product j contains σ_j^2 grams of gold per liter. If we want to get as much gold as possible, what should we do if we have k dollars to spend? The answer is to buy out all the stock of the first k products since they all cost the same, but the first k products have more gold per liter than any of the other products.

The rigorous proof of this claim is omitted here in this note, but only because you were walked through this exact proof on the homework. (It follows from an argument that doing anything else can't be optimal since anything else could be improved by moving it closer to this solution.)

But is *this* assignment of d_j above actually achievable in our original problem of interest? It certainly satisfies all of our inequalities, but that is not sufficient to show it is actually possible — perhaps we could have derived another inequality that makes this assignment impossible. After all, the d_j are defined in (43) from the underlying Q matrix. Maybe no Q matrix exists that satisfies our desired asignments of d_j?

Consequently, the best way to show that it is in fact feasible is to present such a Q. We want the first k rows of Q to all be of unit norm, and the remaining columns to all be 0. Thus, one choice would be to write

$$Q = \begin{bmatrix} I_{k\times k} \\ 0_{(m-k)\times k} \end{bmatrix} = \begin{bmatrix} \vec{e}_1 & \dots & \vec{e}_k \end{bmatrix}. \tag{51}$$

It is straightfoward to verify that the columns of Q form an orthonormal basis of a k-dimensional subspace, so this choice of Q is valid and also shown to be optimal.

So what low-rank approximation does this optimal choice of Q correspond to? Substituting back, we see that

$$\widehat{A} = UXV^\top \tag{52}$$

$$= UQQ^\top \Sigma V^\top. \tag{53}$$

[2] You will see in later courses like 127 and 170 that such problems are actually called "linear programs" and are extremely useful. They arise naturally in lots of settings. In 127 and 170, you will learn ways of thinking about them that are generally useful. In our case here, we can reason about the problem directly.

Now, observe that

$$QQ^\top = \begin{bmatrix} I_{k\times k} \\ 0_{(m-k)\times k} \end{bmatrix} \begin{bmatrix} I_{k\times k} & 0_{k\times(m-k)} \end{bmatrix} = \begin{bmatrix} I_{k\times k} & 0_{k\times(m-k)} \\ 0_{(m-k)\times k} & 0_{(m-k)\times(m-k)} \end{bmatrix} \tag{54}$$

In words, $QQ^\top$ is a diagonal matrix where the first k entries along the diagonal are 1, and the remainder are 0. Thus, we see that

$$\widehat{A} = UQQ^\top \Sigma V^\top \tag{55}$$

$$= U \begin{bmatrix} I_{k\times k} & 0_{k\times(m-k)} \\ 0_{(m-k)\times k} & 0_{(m-k)\times(m-k)} \end{bmatrix} \Sigma V^\top \tag{56}$$

$$= U \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_k & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & \cdots & 0 \end{bmatrix} V^\top \tag{57}$$

$$= \sum_{i=1}^{k} \sigma_i \vec{u}_i \vec{v}_i^\top, \tag{58}$$

as expected. This completes the proof.

The entire proof here is not something that we would expect you to be able to come up with on your own at the level of 16B. However, it is elementary enough that you should be able to follow the steps, and the goal is to help you begin to feel how each part follows naturally.

Contributors:

- Rahul Arya.
- Anant Sahai.
- Ayan Biswas.
- Druv Pai.
- Ashwin Vangipuram.

