Chris Choy
Research Scientist at NVIDIA

Follow

Interesting Properties of Matrix Norms and Singular Values

Matrix norms and singular values have special relationships. Before I forget about them, I'll summarized them in this post.

Definitions

- Schatten p-Norm

  The Schatten p-Norm is defined as the following.[1]

$$\|X\|_{S_p} := \left(\sum_i^n s_i(X)^p \right)^{\frac{1}{p}}$$

- Nuclear Norm

  The nuclear norm of a matrix is defined as a special case of the Schatten p-norm where $p = 1$.

- Frobenius Norm

$$\|X\|_F = \sqrt{\sum_{i=1}^n \sum_{j=1}^n |a_{ij}|^2}$$

- Matrix p-Norm

  Matrix p-norm is defined as

$$\|A\|_p = \sup_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_p}$$

  In another word, matrix p-Norm is defined as the largest scalar that you can get for a unit vector e.

- Spectral Norm

  Largest singular value of a matrix $\sigma_1(X)$.

  Special case of the matrix p-norm where $p = 2$ when the matrix X is positive semi-definite. For negative definite matrix, the matrix 2-norm is not necessarily the largest norm.

Lemmas

1. $A \in \mathbf{S}^n$ $tr(A) = \sum_i^n \lambda_i = \|A\|_{S_1}$

   Trace of a symmetric matrix A is equal to the sum of eigen values. Let A be a symmetric matrix $A \in \mathcal{S}^n$. Then there exists a orthogonal matrix U and diagonal matrix Λ such that $A = U \Lambda U^T$.

$$
\begin{aligned}
tr(A) &= tr(A^T) && (1)\\
&= \sum_i^n e_i^T A^T e_i && (2)\\
&= \sum_i^n e_i^T U^T \Lambda U e_i && (3)\\
&= \sum_i^n \sum_j^n u_{ji}^T \lambda_j u_{ji} && (4)\\
&= \sum_j^n \lambda_j \sum_i uji^2 && (5)\\
&= \sum_j^n \lambda_j && (6)
\end{aligned}
$$

Where $U = [u_1, u_2, u_3, \dots u_n]$. We used the fact that $u_i^T u_i = 1$.

2. $A \in \mathbf{S}_+^n$ $tr(A) = \sum_i^n |\sigma_i| = |A|_{S_1}$

   Trace of a positive semi-definite matrix A is equal to the L1 norm of singular values, or is equal to the Schatten 1-Norm (Nuclear Norm).

   This is the direct extension of Lemma 1.

$$
\begin{aligned}
tr(A) &= \sum_i^n \lambda_i && (7)\\
&= \sum_i^n \sigma_i && (8)\\
&= \sum_i^n \|\sigma_i\| && (9)\\
&= \|A\|_{S_1} && (10)
\end{aligned}
$$

   Since the L1 norm of singular values enforce sparsity on the matrix rank, yhe result is used in many application such as low-rank matrix completion and matrix approximation.

3. $\|X\|_F = \sqrt{\sum_i^n \sigma_i^2} = \|X\|_{S_2}$

   Frobenius norm of a matrix is equal to L2 norm of singular values, or is equal to the Schatten 2 norm.

$$
\begin{aligned}
\|X\|_F &= \sqrt{tr(X^T X)} && (11)\\
&= \sqrt{tr(V\Sigma U^T U \Sigma V^T)} && (12)\\
&= \sqrt{tr(V\Sigma^2 V^T)} && (13)\\
&= \sqrt{\sum_i \sigma_i^2} && (14)\\
&= \|X\|_{S_2} && (15)
\end{aligned}
$$

4. $\|A\|_1 = \max_j \sum_i^n |a_{ij}|$

L1 matrix norm of a matrix is equal to the maximum of L1 norm of a column of the matrix.

To begin with, the solution of L1 optimization usually occurs at the corner. If the function of interest is piece-wise linear, the extrema always occur at the corners. One can show this by showing that $\frac{|Ax|_1}{|x|_1}$ is Lipschitz except $x = 0$ and differentiate $\frac{|Ax|_1}{|x|_1}$ w.r.t. x_i. Then show that they are non-zero. Use the fact on line 1 and 2.

Finally, we can just compute the L1 norms of each column. Let's reformulate the problem.

$$\begin{aligned} \|A\|_1 &= \sup_{x \neq 0} \frac{\|Ax\|_1}{\|x\|_1} && (16) \\ &= \max_{\|x\|_1 = 1} \sum_j \sum_i^n |a_{ij}||x_j| && (17) \\ &= \max_j \sum_i |a_{ij}| && (18) \end{aligned}$$

5. $\|A\|_\infty = \max_i \sum_j^n |a_{ij}|$

   The supremum occurs at the corner of the hypercube since infinity nomr of a vector is the absolute value of the largest element in it. Thus,

$$\begin{aligned} \|A\|_\infty &= \max \frac{\|Ax\|_\infty}{\|x\|_\infty} && (19) \\ &= \max_{\|x_j\| = 1} \|Ax\|_\infty && (20) \\ &= \max_i \|a_i^T\| && (21) \end{aligned}$$

   Where a_i^T is the i th row of the matrix A.

1. http://en.wikipedia.org/wiki/Schatten_norm ↩

Tags: matrix norm

Categories: research

Created: January 13, 2015
Updated: January 14, 2015

Previous | Next

LEAVE A COMMENT