Alex Lemon
EE263 (Introduction to Linear Dynamical Systems)
Problem of the Day 13
August 5, 2014

1 Representations of ellipsoids

Consider an ellipsoid $\mathcal{E} \subseteq \mathbb{R}^n$ that is centered at the origin, and has nonzero volume. We have seen two representations for such an ellipsoid:

$$\mathcal{E}_1(S) = \{x \in \mathbb{R}^n : x^\mathsf{T} S x \leq 1\} \qquad \text{and} \qquad \mathcal{E}_2(A) = \{Az : \|z\| \leq 1\},$$

where $S \in \mathbb{R}^{n\times n}$ is symmetric and positive definite, and $A \in \mathbb{R}^{n\times n}$ is nonsingular.

(a) Given a nonsingular matrix $A \in \mathbb{R}^{n\times n}$, explain how to find a symmetric, positive-definite matrix $S \in \mathbb{R}^{n\times n}$ such that $\mathcal{E}_1(S) = \mathcal{E}_2(A)$. Is there a unique S corresponding to a given A? If so, explain why; if not, explain why not, and how to find all matrices S corresponding to a given A.

(b) Given a symmetric, positive-definite matrix $S \in \mathbb{R}^{n\times n}$, explain how to find a nonsingular matrix $A \in \mathbb{R}^{n\times n}$ such that $\mathcal{E}_1(S) = \mathcal{E}_2(A)$. Is there a unique A corresponding to a given S? If so, explain why; if not, explain why not, and how to find all matrices A corresponding to a given S.

2 Confidence ellipsoids

Consider a measurement model

$$y = Ax + v,$$

where $y \in \mathbb{R}^m$ is a measurement (which is known), $x \in \mathbb{R}^n$ is a vector of unknown parameters (which we want to estimate), $v \in \mathbb{R}^m$ is measurement noise (which is unknown), and $A \in \mathbb{R}^{m\times n}$ describes the measurement system. We will assume that A is skinny and full rank, and v is norm-bounded: that is, there is a known $\alpha > 0$ such that $\|v\| \leq \alpha$.

(a) First, consider the specific case when

$$A = \begin{bmatrix} 2 \\ -1 \end{bmatrix}, \qquad y = \begin{bmatrix} 9 \\ 23 \end{bmatrix}, \qquad \text{and} \qquad \alpha = 25.$$

We say that the parameter $x \in \mathbb{R}$ is consistent with the model if there exists a noise vector $v \in \mathbb{R}^m$ such that $y = Ax + v$, and $\|v\| \leq \alpha$. Find the set of all parameters x that are consistent with the specific model given above. You can do this by finding the set of all x such that $\|y - Ax\| \leq \alpha$.

(b) Your friend proposes a different method for determining the set of consistent parameters. Since $y = Ax + v$, and $A^\dagger A = I$ for a skinny and full rank matrix, he argues that

$$\hat{x} = A^\dagger y = A^\dagger(Ax + v) = x + A^\dagger v,$$

and hence that

$$x = \hat{x} - A^\dagger v,$$

where $\hat{x} = A^\dagger y$ is the least-squares estimate of x. Therefore, the set of all parameter vectors x that are consistent with the model is

$$\hat{x} - A^\dagger\{v \in \mathbb{R}^m : \|v\| \leq \alpha\}.$$

(i) According to your friend's analysis, what is the set of all parameter vectors x that are consistent with the specific model given in (a)? Is the consistent with the result that you derived earlier?

(ii) If your friend's analysis does not produce the same answer you derived earlier, why is there a difference? Which do you think is correct: your friend's analysis, or the answer you derived earlier?

(c) Now consider the general case. Let the singular-value decomposition of A be $A = U\Sigma V^\mathsf{T}$. By "completing the square," show that the set of all parameters $x \in \mathbb{R}^n$ that are consistent with the model is

$$\left\{\hat{x} + V\Sigma^{-1}z : \|z\| \leq \sqrt{\alpha^2 - \|\hat{\epsilon}\|^2}\right\}.$$

(d) The file `confidence_ellipsoids_data.m` contains values of A, y, and α. For this example, the vector of parameters has dimension $n = 2$. Plot the set of all parameter vectors x that are consistent with the model.

3 The Frobenius norm

The Frobenius norm of a matrix $A \in \mathbb{R}^{m\times n}$ is defined to be

$$\|A\|_F = \sqrt{\operatorname{tr}(A^\mathsf{T} A)}.$$

(a) Show that

$$\|A\|_F = \left(\sum_{i=1}^m \sum_{j=1}^n |a_{ij}|^2\right)^{\frac{1}{2}}.$$

Thus, the Frobenius norm is simply the Euclidean norm of a matrix, when we think of the matrix as an element of $\mathbb{R}^{mn}$. Additionally, note that the Frobenius norm is much easier to compute than the spectral norm.

(b) Show that if $U \in \mathbb{R}^{m\times m}$ and $V \in \mathbb{R}^{n\times n}$ are orthogonal, then

$$\|UA\|_F = \|AV\|_F = \|A\|_F.$$

Thus, multiplication by orthogonal matrices on the left or right does not change the Frobenius norm.

(c) Show that

$$\|A\|_F = \sqrt{\sigma_1^2 + \cdots + \sigma_r^2},$$

where $\sigma_1, \ldots, \sigma_r$ are the nonzero singular values of A. Use this result to deduce that

$$\sigma_{\max}(A) \leq \|A\|_F \leq \sqrt{r}\,\sigma_{\max}(A).$$

In particular, we have that $\|Ax\| \leq \|A\|_F\|x\|$ for all $x \in \mathbb{R}^n$.