EE562a Elliptical Region Bound – Keith M. Chugg

Second Moment Bounds

The standard bounds from probability (i.e., Chebychev, Markov, and variations) can be used to bound the probability that the random vector $\mathbf{x}(u)$ is outside some region of the plan. Direct application leads to

$$\Pr\{\|\mathbf{x}(u)\| \geq \epsilon\} = \Pr\left\{\|\mathbf{x}(u)\|^2 \geq \epsilon^2\right\} \tag{1}$$

$$\leq \frac{1}{\epsilon^2}\mathbb{E}\left\{\|\mathbf{x}(u)\|^2\right\} = \frac{1}{\epsilon^2}\mathrm{tr}\left(\mathbf{R}_{\mathbf{x}}\right). \tag{2}$$

We have used the simple fact that

$$\mathbb{E}\left\{\|\mathbf{x}(u)\|^2\right\} = \mathbb{E}\left\{\mathrm{tr}\left(\mathbf{x}(u)\mathbf{x}^{\dagger}(u)\right)\right\} = \mathrm{tr}\left(\mathbb{E}\left\{\mathbf{x}(u)\mathbf{x}^{\dagger}(u)\right\}\right) = \mathrm{tr}\left(\mathbf{R}_{\mathbf{x}}\right). \tag{3}$$

This also leads directly to a centered version

$$\Pr\{\|\mathbf{x}(u) - \mathbf{m}_{\mathbf{x}}\| \geq \epsilon\} \leq \frac{1}{\epsilon^2}\mathrm{tr}\left(\mathbf{K}_{\mathbf{x}}\right). \tag{4}$$

If $n = 2$ and $\mathbf{x}(u)$ is real, the above relations bound the probability that $\mathbf{x}(u)$ is outside a circle of radius ϵ. A more "efficient" bound may be obtained by taking into account the directional preference of $\mathbf{x}(u)$. Specifically, assume that $\mathbf{K}_{\mathbf{x}}$ is non-singular and has been factored into $\mathbf{K}_{\mathbf{x}} = \mathbf{H}\mathbf{H}^{\dagger}$. Then it makes sense to start the bound using the whitened version of $\mathbf{x}(u)$, namely $\mathbf{w}(u) = \mathbf{H}^{-1}\mathbf{x}_0(u)$, which has no directional preference (i.e., $\mathbf{K}_{\mathbf{w}} = \mathbf{I}$). For this white random vector the circular region bound is intuitively reasonable

$$\Pr\left\{\|\mathbf{w}(u)\|^2 \geq \epsilon^2\right\} \leq \frac{1}{\epsilon^2}\mathrm{tr}\left(\mathbf{K}_{\mathbf{w}}\right) = \frac{n}{\epsilon^2}. \tag{5}$$

Noting that

$$\|\mathbf{w}(u)\|^2 = (\mathbf{H}^{-1}\mathbf{x}_0(u))^{\dagger}\mathbf{H}^{-1}\mathbf{x}_0(u) = \mathbf{x}_0^{\dagger}(u)\mathbf{H}^{-\dagger}\mathbf{H}^{-1}\mathbf{x}_0(u) = \mathbf{x}_0^{\dagger}(u)\mathbf{K}_{\mathbf{x}}^{-1}\mathbf{x}_0(u), \tag{6}$$

yields the final form of the "Elliptical Region Bound"

$$\Pr\left\{(\mathbf{x}(u) - \mathbf{m}_{\mathbf{x}})^{\dagger}\mathbf{K}_{\mathbf{x}}^{-1}(\mathbf{x}(u) - \mathbf{m}_{\mathbf{x}}) \geq \epsilon^2\right\} \leq \frac{n}{\epsilon^2}. \tag{7}$$

As the name implies, this bounds the probability that $\mathbf{x}(u)$ is outside of an elliptical region of the plane (i.e., for $n = 2$ and real $\mathbf{x}(u)$). We will show this shortly, but first consider another use for the bound – we can specify, with a lower bound on the probability, that $\mathbf{x}(u)$ is inside the elliptical region.

Probability Regions

The standard Chebychev bound for random variables can be used to derive the "3–σ rule-of-thumb." Specifically, for a real random variable $z(u)$, the Chebychev bound is

$$\Pr\{|z(u) - m_z| \geq \epsilon\} = \Pr\left\{|z(u) - m_z|^2 \geq \epsilon^2\right\} \leq \frac{1}{\epsilon^2}\mathbb{E}\left\{|z(u) - m_z|^2\right\} = \frac{\sigma_z^2}{\epsilon^2}. \tag{8}$$

Suppose that we want a "90% probability region"[1] for $z(u)$ – we would like to specify an interval of the real line, with the probability that $z(u)$ is in this interval ≥ 0.9. This is obtained by choosing $\epsilon^2 = 0.1\sigma_z^2$ in the Chebychev bound, which results in

$$\text{Pr}\left\{|z(u) - m_z| \geq \sqrt{10}\sigma_z\right\} \leq 0.1, \tag{9}$$

or the complement statement

$$\text{Pr}\left\{|z(u) - m_z| \leq \sqrt{10}\sigma_z\right\} > 0.9. \tag{10}$$

Approximating $\sqrt{10}$ by 3, the rule-of-thumb is that *any* random variable is within 3σ of it's mean with 90% probability.

We can generalize this 90% probability region for random vectors using the Elliptical Region Bound – just choose $\epsilon^2 = 10n$ which implies

$$\text{Pr}\left\{\mathbf{x}(u) \in \mathcal{A}\right\} > 0.9 \quad \mathcal{A} = \{\mathbf{v} : (\mathbf{v} - \mathbf{m_x})^\dagger \mathbf{K}_\mathbf{x}^{-1} (\mathbf{v} - \mathbf{m_x}) < 10n\}. \tag{11}$$

In other words $\mathcal{A}$ is a 90% probability region for $\mathbf{x}(u)$.

The $n = 2$ Special Case Since the probability region will be used as a tool to build intuition, we will concentrate on the case where $n = 2$ and $\mathbf{x}(u)$ is real. In this case, the probability region is the subset of the plane defined by

$$\mathcal{A} = \{\mathbf{v} \in \mathcal{R}^2 : (\mathbf{v} - \mathbf{m_x})^{\mathrm{t}} \mathbf{K}_\mathbf{x}^{-1} (\mathbf{v} - \mathbf{m_x}) < 20\}. \tag{12}$$

This region is an ellipse with principle axes aligned with the eigen directions (see the next section for a proof). The length of the major and minor axes are proportional to the square-root of the maximum and minimum eigenvalues, respectively. The exact shape and dimension are diagrammed in Fig. 1

If realizations of $\mathbf{x}(u)$ were generated, they would all lie inside this elliptical region with probability greater than 0.9; for this reason it is also useful to think of this region as a rough "scatter plot" of the values of $\mathbf{x}(u)$. When we get to Gaussian random vectors, we will see that this is especially useful.

The Region Really is an Ellipse

Let's show the details of the claim that the 2-dimension regions are ellipses. Specifically, the following is true:

Theorem *If* $\mathbf{K}$ is a real (2×2) positive definite, symmetric matrix, *then* the equation

$$\mathbf{v}^{\mathrm{t}} \mathbf{K}^{-1} \mathbf{v} = C^2 \tag{13}$$

defines an ellipse with principle axes aligned with the orthonormal eigenvectors, and proportional to the corresponding eigenvalues.

[1] There is nothing magical about 90%, just that it is a nice round figure. You can modify the development to get any probability level you like.

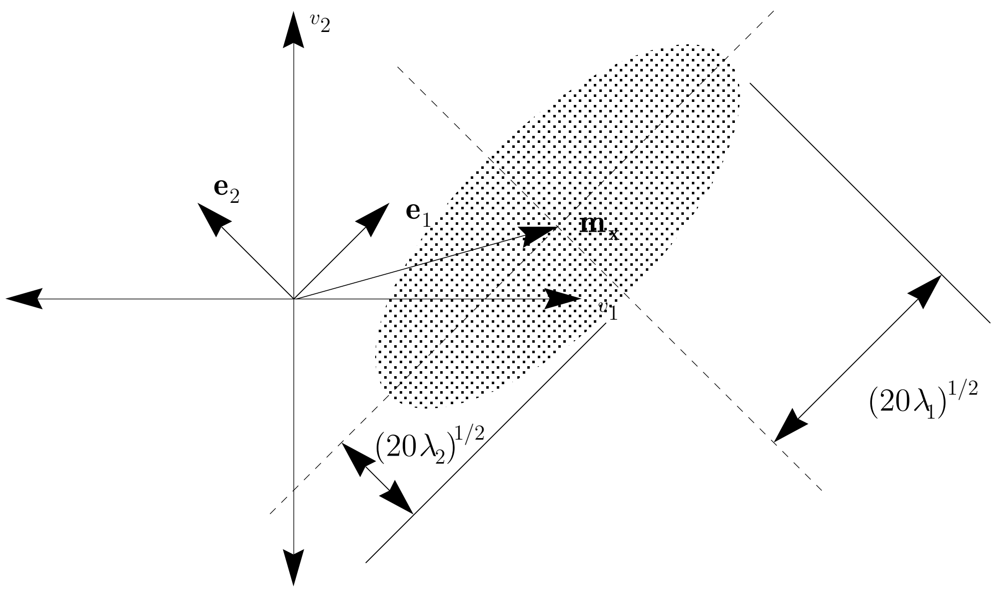

Figure 1: The region containing at least 90% probability mass for $\mathbf{x}(u)$.

Proof The proof is obtained by a change of coordinates. Let $\mathbf{w} = \mathbf{E}^{\mathrm{t}}\mathbf{v}$ where $\mathbf{K} = \mathbf{E}\boldsymbol{\Lambda}\mathbf{E}^{\mathrm{t}}$ is the symmetry transform defined by the orthonormal e-vectors, then in terms of $\mathbf{w}$, the relation becomes

$$\begin{aligned} C^2 &= (\mathbf{E}\mathbf{w})^{\mathrm{t}}\mathbf{K}^{-1}\mathbf{E}\mathbf{w} && (14)\\ &= \mathbf{w}^{\mathrm{t}}\mathbf{E}^{\mathrm{t}}(\mathbf{E}\boldsymbol{\Lambda}^{-1}\mathbf{E}^{\mathrm{t}})\mathbf{E}\mathbf{w} && (15)\\ &= \mathbf{w}^{\mathrm{t}}\boldsymbol{\Lambda}^{-1}\mathbf{w} && (16)\\ &= \frac{w_1^2}{\lambda_1} + \frac{w_2^2}{\lambda_2}. && (17) \end{aligned}$$

The last line defines an ellipse in the (w_1, w_2) coordinate system which crosses the w_1 axis at $C\sqrt{\lambda_1}$ and w_2 axis at $C\sqrt{\lambda_2}$. These principle axes are aligned with the eigenvectors since the component of $\mathbf{w}$ in the $\mathbf{e}_i$ direction is simply w_i. Since multiplication by an orthogonal matrix (e.g., $\mathbf{E}^{\mathrm{t}}$) corresponds to rotation and/or reflection in the plane, the region defined by this equation is always as sketched in Fig. 2

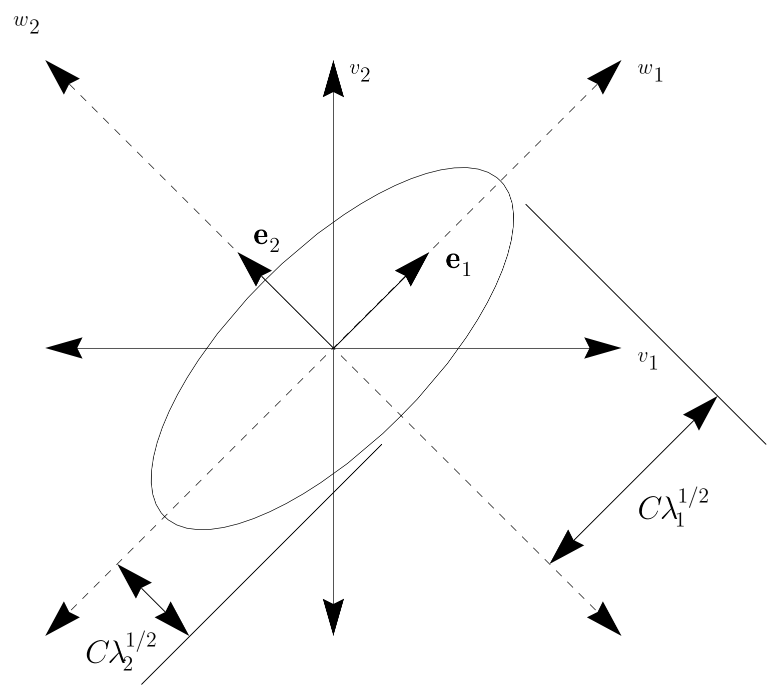

Figure 2: The probability of being outside this ellipse is less than $2/C^2$.