Stuart & Sun, "Matrix Perturbation Thy".

3.5. The Field of Values

The quadratic form $x^{\mathrm{H}}Ax$ plays an important role in many applications. This subsection is devoted to the values that such a form can attain for a given matrix. We begin with a definition.

Definition 3.10. *Let $A \in \mathbf{C}^{n\times n}$. The set*

$$\mathcal{F}(A) \stackrel{\text{def}}{=} \{x^{\mathrm{H}}Ax : \|x\|_2 = 1\}$$

is called the FIELD OF VALUES *of A.*

The set $\mathcal{F}(A)$ is bounded and closed in $\mathbf{C}$. From Definition 3.10, it is easy to verify the following properties of $\mathcal{F}(A)$:

1. $\mathcal{F}(\alpha A + \beta I) = \alpha\mathcal{F}(A) + \beta, \quad \alpha, \beta \in \mathbf{C}$;
2. $\mathcal{L}(A) \subseteq \mathcal{F}(A)$;
3. If U is unitary, $\mathcal{F}(U^{\mathrm{H}}AU) = \mathcal{F}(A)$;
4. $\mathcal{F}(A+B) \subseteq \mathcal{F}(A) + \mathcal{F}(B)$.

An important and far-reaching result is that the field of values of a matrix is convex; that is, it contains any line whose endpoints lie in it. Nothing in the proof of the following theorem is used later, and it may be skipped without loss.

Theorem 3.11 (Toeplitz–Hausdorff). *The field of values $\mathcal{F}(A)$ is a convex set.*

Proof. Let $\rho, \sigma \in \mathcal{F}(A)$. Since $\mathcal{F}(\alpha A+\beta I) = \alpha\mathcal{F}(A)+\beta$ and $\alpha\mathcal{F}(A)+\beta$ is convex if and only if $\mathcal{F}(A)$ is convex, we may assume without loss of generality that $\rho = 0$ and $\sigma = 1$. Thus there are vectors x_0 and x_1 of 2-norm one such that

$$x_0^{\mathrm{H}}Ax_0 = 0 \text{ and } x_1^{\mathrm{H}}Ax_1 = 1,$$

which implies that x_0 and x_1 are linearly independent. By multiplying x_0 by a scalar of absolute value one we may further assume that

$$\Re\, x_0^{\mathrm{H}}x_1 = 0.$$

We must show that any $\tau \in [0,1]$ is in the field of values of A. By the linear independence of x_0 and x_1, the vector $(1-\lambda)x_0 + \lambda x_1$ is nonzero. Consequently the function

$$\varphi(\lambda) = \frac{[(1-\lambda)x_0 + \lambda x_1]^{\mathrm{H}} A [(1-\lambda)x_0 + \lambda x_1]}{\|(1-\lambda)x_0 + \lambda x_1\|_2^2} = \frac{|\lambda|^2}{\|(1-\lambda)x_0 + \lambda x_1\|_2^2}$$

is continuous and real when λ is real. Moreover, $\varphi(\lambda) \in \mathcal{F}(A)$ for all λ. Since $\varphi(0) = 0$ and $\varphi(1) = 1$, by the intermediate value theorem there is a $\lambda \in [0,1]$ such that $\tau = \varphi(\lambda) \in \mathcal{F}(A)$. ∎

The field of values of a matrix A is closely related to the eigenvalues of A. In particular if $Ax = \lambda x$ with $\|x\|_2 = 1$, then $\lambda = x^{\mathrm{H}} A x \in \mathcal{F}(A)$. Since the field of values is convex, $\mathcal{F}(A)$ must contain the smallest convex set containing all the eigenvalues of A, that is, the set

$$\mathcal{H}[\mathcal{L}(A)] = \left\{ \sum_{\lambda_i \in \mathcal{L}(A)} \theta_i \lambda_i : \theta_i \geq 0, \sum \theta_i = 1 \right\} \tag{3.7}$$

($\mathcal{H}[\mathcal{L}(A)]$ is called the CONVEX HULL of $\mathcal{L}(A)$). Unfortunately, the field of values can be bigger than the convex hull of $\mathcal{L}(A)$, as the following example shows.

Example 3.12. *Let*

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}.$$

Then $\mathcal{H}[\mathcal{L}(A)] = \{0\}$. *But* $\mathcal{F}(A) = \{z \in \mathbf{C} : |z| \leq \frac{1}{2}\}$.

There is one important class of matrices for which the field of values coincides with the convex hull of its eigenvalues.

Theorem 3.13. *If A is normal, then*

$$\mathcal{F}(A) = \mathcal{H}[\mathcal{L}(A)]. \tag{3.8}$$

Proof. Since the normal matrix A has a decomposition $A = U^{\mathrm{H}} \Lambda U$, where U is unitary and $\Lambda = \mathrm{diag}(\lambda_1, \ldots, \lambda_n)$, we have

$$\begin{aligned}
\mathcal{F}(A) &= \{(Ux)^{\mathrm{H}} \Lambda (Ux) : x \in \mathbf{C}^n, \|x\|_2 = 1\} \\
&= \{y^{\mathrm{H}} \Lambda y : y \in \mathbf{C}^n, \|y\|_2 = 1\} \\
&= \left\{\textstyle\sum_{i=1}^n |y_i|^2 \lambda_i : \sum_i |y_i|^2 = 1\right\}.
\end{aligned}$$

Since $\sum |y_i|^2 = 1$, this last equation is clearly equivalent to (3.7). ∎

Theorem (Töplitz-Hausdorff): The field of values, $\mathcal{F}(A)$, with $A \in \mathbb{C}^{n\times n}$, is a convex set.

▶ Suppose $\rho, \sigma \in \mathcal{F}(A)$. We must show that for $t \in [0,1]$, $\tau(t) = (1-t)\rho + t\sigma \in \mathcal{F}(A)$.

Since $\rho, \sigma \in \mathcal{F}(A) \Rightarrow \exists x_0, x_1 \in \mathbb{C}^n : x_0^H A x_0 = \rho$, $x_1^H A x_1 = \sigma$, and we must find $z = z(t)$ so that $z^H A z = \tau(t)$, $t \in [0,1]$.

Taking advantage of the property $\mathcal{F}(aA + \beta I) = a\mathcal{F}(A) + \beta$, we choose a, β so that $a\rho + b = 0$, $a\sigma + b = 1$, i.e.

$$a = 1/(\sigma - \rho), \quad \beta = -\rho/(\sigma-\rho).$$

Then, if $B = aA + \beta I$, $0, 1 \in \mathcal{F}(B)$ with $x_0^H B x_0 = 0$, $x_1^H B x_1 = 1$ and we must show that $\forall t \in [0,1]$, we can find $z : z^H B z = t$.

(recall that $\|x_0\| = \|x_1\| = 1$ in the defn. for $\mathcal{F}(A)$ etc.).

<u>Note</u> that, $\forall x$, $x^H B x = (e^{i\vartheta} x)^H B (e^{i\vartheta} x)$, which gives us freedom to affect things later.

Now, let $z = (1-\lambda)x_0 + \lambda x_1$; $\lambda \in [0,1]$, and consider

$$\frac{z^H B z}{z^H z} = \left((1-\lambda)x_0^H + \lambda x_1^H\right) B \left((1-\lambda)x_0 + \lambda x_1\right) / \|z\|^2$$

$$= \frac{\lambda^2}{\|z\|^2} + \frac{\lambda(1-\lambda)}{\|z\|^2}\left\{x_0^H B x_1 + x_1^H B x_0\right\}$$

Since the second term vanishes for $\lambda = 0, 1$, both terms are continuous and the sum equals 0 at $\lambda = 0$ and 1 at $\lambda = 1$, to finish the proof we must show that the quantity in $\{\cdots\}$ can be chosen to be real

$\twoheadrightarrow$

Now, $B = B_H + B_S$ where $B_H = \frac{1}{2}(B + B^H)$ is Hermitian and $B_S = \frac{1}{2}(B^H - B^S)$ is skew Hermitian. We have $xBy = x\cancel{B^{\#}_H}y + xB_S y$.

$$\text{(i)}\quad x_0^H B_H x_1 = \langle x_0, B_H x_1\rangle = \langle B_H x_1, x_0\rangle^*$$
$$= (x_1^H B_H^H x_0)^* = (x_1^H B_H x_0)^*$$

$(x_0^H B_H x_1) + (x_0^H B_H x_1)^* = x_0^H B_H x_1 + x_1^H B_H x_0 = 2\mathrm{Re}$

$$\text{So}\quad x_0^H B_H x_1 + x_1^H B_H x_0 = 2\mathrm{Re}\, x_0^H B_H x_1$$

$$\text{(ii)}\quad x_0^H B_S x_1 = \langle x_0, B_S x_1\rangle = \langle B_S x_1, x_0\rangle^*$$
$$= (x_1^H B_S^H x_0)^* = -(x_1^H B_S x_0)^*$$

$(x_0^H B_S x_1) - (x_0^H B_S x_1)^* = x_0^H B_S x_1 +$

$\quad\quad \| $

$2i\,\mathrm{Im}(x_0^H B_S x_1)$

$$\text{So}\quad x_0^H B_S x_1 + x_1^H B_S x_0 = 2i\,\mathrm{Im}\, x_0^H B_S x_1$$

~~But then~~ $\cancel{\{x_0^H B x_1 + x_1^H B x_0\} \neq 2 x_0^H B_H x_1}$.

To show that we can make this term ~~real~~ vanish, let

$y = B_H x_1$, then

$$x_0^H y = (x_{0,1}^T - i x_{0,2}^T)(y_1 + i y_2)$$
$$= (\langle x_{0,1}, y_1\rangle + \langle x_{0,2}, y_2\rangle) + i\underbrace{(\langle x_{0,1}, y_2\rangle - \langle x_{0,2}, y_1\rangle)}$$

$$u + iv$$

while $(e^{i\theta}x_0)^H y = (\cos\theta - i\sin\theta)\, x_0^H y$

$$= (\cos\theta - i\sin\theta)(u + iv)$$
$$= (\cos\theta\, u + \sin\theta\, v) + i(\sin\theta\, u - \cos\theta\, v)$$

The choice ~~of~~ $\theta = \tan^{-1}(u/v)$ makes the imaginary part vanish (unless $u = v = 0$, but then the entire term vanishes and there is nothing to do!). This completes the proof.

Since: $x_0 x_1 = x_1 x_0 \Rightarrow x_0 B x_1 + x_1 B x_0$