NOTES RELATED TO SVD DERIVATION

Given $\underline{\mathbf{f}} = \underline{\underline{\mathbf{M}}}\,\underline{\mathbf{i}}$ and a growth in (length)^2 of $s \equiv \dfrac{|\underline{\mathbf{f}}|^2}{|\underline{\mathbf{i}}|^2} = \dfrac{\underline{\mathbf{f}}\cdot\underline{\mathbf{f}}}{\underline{\mathbf{i}}\cdot\underline{\mathbf{i}}} = \dfrac{\underline{\mathbf{i}}^{\mathrm{T}}\,\underline{\underline{\mathbf{M}}}^{\mathrm{T}}\,\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{i}}}{\underline{\mathbf{i}}^{\mathrm{T}}\,\underline{\mathbf{i}}}$ (note $\underline{\mathbf{i}}\cdot\underline{\mathbf{i}} \equiv \underline{\mathbf{i}}^{\mathrm{T}}\,\underline{\mathbf{i}}$), what vector $\underline{\mathbf{i}}$ will extremize s ?

Let $\underline{\mathbf{I}}_n$ be some orthonormal basis. We may as well let $\underline{\mathbf{i}}$ have unit length ($\underline{\mathbf{i}}\cdot\underline{\mathbf{i}} = 1$). If we work in 2 dimensions we can represent $\underline{\mathbf{i}}$ in the $\underline{\mathbf{I}}_n$ basis as $\underline{\mathbf{i}} = \cos(q)\,\underline{\mathbf{I}}_1 + \sin(q)\,\underline{\mathbf{I}}_2$. Substituting this into the definition of s gives,

$$s = \frac{\underline{\mathbf{f}}\cdot\underline{\mathbf{f}}}{\underline{\mathbf{i}}\cdot\underline{\mathbf{i}}} = \underline{\mathbf{f}}\cdot\underline{\mathbf{f}} = \left(\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{i}}\right)\cdot\left(\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{i}}\right) = \left(\cos(q)\,\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_1 + \sin(q)\,\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_2\right)\cdot\left(\cos(q)\,\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_1 + \sin(q)\,\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_2\right)$$

$$= \cos^2(q)\left|\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_1\right|^2 + \sin^2(q)\left|\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_2\right|^2 + 2\cos(q)\sin(q)\left(\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_1\right)\cdot\left(\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_2\right)$$

To find the vectors $\underline{\mathbf{i}} = \cos(q)\,\underline{\mathbf{I}}_1 + \sin(q)\,\underline{\mathbf{I}}_2$ that extremize s , let

$$0 = \frac{\mathrm{d}s}{\mathrm{d}q} = \sin(2q)\left(-\left|\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_1\right|^2 + \left|\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_2\right|^2\right) + 2\cos(2q)\left(\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_1\right)\cdot\left(\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_2\right)$$

We want to find all the vectors $\underline{\mathbf{i}}$ that satisfy the above equality. We can simply this task by choosing a convenient basis (this won't limit the possible values of $\underline{\mathbf{i}}$). Consider the basis $\underline{\mathbf{I}}_n$ that satisfies

$$0 = \left(\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_1\right)\cdot\left(\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_2\right) = \underline{\mathbf{I}}_1^{\mathrm{T}}\,\underline{\underline{\mathbf{M}}}^{\mathrm{T}}\,\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_2 = \underline{\mathbf{I}}_1\cdot\left(\underline{\underline{\mathbf{M}}}^{\mathrm{T}}\,\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_2\right)$$

For this to be true, $\left(\underline{\underline{\mathbf{M}}}^{\mathrm{T}}\,\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_2\right)$ must be perpendicular to $\underline{\mathbf{I}}_1$. Since $\underline{\mathbf{I}}_2$ is also perpendicular to $\underline{\mathbf{I}}_1$, $\left(\underline{\underline{\mathbf{M}}}^{\mathrm{T}}\,\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_2\right)$ must be parallel to $\underline{\mathbf{I}}_2$,

$$\left(\underline{\underline{\mathbf{M}}}^{\mathrm{T}}\,\underline{\underline{\mathbf{M}}}\right)\underline{\mathbf{I}}_2 = l\,\underline{\mathbf{I}}_2$$

$\underline{\mathbf{I}}_2$, then, is an eigenvector of $\left(\underline{\underline{\mathbf{M}}}^{\mathrm{T}}\,\underline{\underline{\mathbf{M}}}\right)$. We should note that in this basis $\underline{\mathbf{I}}_1$ also solves the above eigenvalue equation, and that only these 2 linearly independent eigenvectors exist for the 2x2 matrix $\left(\underline{\underline{\mathbf{M}}}^{\mathrm{T}}\,\underline{\underline{\mathbf{M}}}\right)$.

In this basis, the vectors $\underline{\mathbf{i}}$ that extremize s satisfy

$$0 = \frac{\mathrm{d}s}{\mathrm{d}q} = \sin(2q)\left(-\left|\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_1\right|^2 + \left|\underline{\underline{\mathbf{M}}}\,\underline{\mathbf{I}}_2\right|^2\right)$$

Which is only true when $q = k\dfrac{p}{2},\quad k = 0,1,2,\ldots$ (i.e., when $\underline{\mathbf{i}}$ is aligned with one of the two basis vectors).

Thus $\underline{\mathbf{i}} = \underline{\mathbf{I}}_1$ and $\underline{\mathbf{i}} = \underline{\mathbf{I}}_2$ are the two vectors that extremize s (one grows most, one least). The value of l can be obtained by plugging $\underline{\mathbf{I}}_2 = \underline{\mathbf{i}}$ into the above eigenvalue equation and dotting $\underline{\mathbf{i}}$ into both sides,

$$\underline{\mathbf{i}}\cdot\left(\underline{\underline{\mathbf{M}}}^{\mathrm{T}}\,\underline{\underline{\mathbf{M}}}\right)\underline{\mathbf{i}} = l\,\underline{\mathbf{i}}\cdot\underline{\mathbf{i}}$$

$$\underline{\mathbf{f}}\cdot\underline{\mathbf{f}} = l$$

Since we know $\underline{\mathbf{f}}\cdot\underline{\mathbf{f}} = s$, the eigenvalues are $l = s$.

In summary, the growth in length of a vector $\underline{\mathbf{i}}$ under operation by $\underline{\underline{\mathbf{M}}}$ will be extremized when $\underline{\mathbf{i}}$ satisfies

$$\left(\underline{\underline{\mathbf{M}}}^{\mathrm{T}}\,\underline{\underline{\mathbf{M}}}\right)\underline{\mathbf{i}} = s\,\underline{\mathbf{i}}$$

where s represents the change in absolute square of $\underline{\mathbf{i}}$ after operation by $\underline{\underline{\mathbf{M}}}$. The vectors $\underline{\mathbf{i}}$ are the initial singular vectors, and $\sqrt{s}$ represents the singular values. Operating on the above equality from the left with $\underline{\underline{\mathbf{M}}}$ gives

$$\left(\underline{\underline{\mathbf{M}}}\,\underline{\underline{\mathbf{M}}}^{\mathrm{T}}\right)\underline{\mathbf{f}} = s\,\underline{\mathbf{f}}$$

The vectors $\underline{\mathbf{f}}$ that satisfy this equality are the final time singular vectors.