Similarly, $\delta(\mathcal{N},\mathcal{M}) = \|(\mathbf{I}-\mathbf{P}_{\mathcal{M}})\mathbf{P}_{\mathcal{N}}\|_2 = \|\mathbf{P}_{\mathcal{N}}(\mathbf{I}-\mathbf{P}_{\mathcal{M}})\|_2$. Let $\mathbf{U} = \left(\mathbf{U}_1 \,|\, \mathbf{U}_2\right)$ and $\mathbf{V} = \left(\mathbf{V}_1 \,|\, \mathbf{V}_2\right)$ be orthogonal matrices such that

$$\begin{array}{ll} R(\mathbf{U}_1) = \mathcal{M} & R(\mathbf{U}_2) = \mathcal{M}^{\perp} \\ R(\mathbf{V}_1) = \mathcal{N}^{\perp} & R(\mathbf{V}_2) = \mathcal{N} \end{array} \tag{5.13.11}$$

so that $\mathbf{P}_{\mathcal{M}} = \mathbf{U}_1\mathbf{U}_1^T$ and $\mathbf{I}-\mathbf{P}_{\mathcal{N}} = \mathbf{P}_{\mathcal{N}^{\perp}} = \mathbf{V}_1\mathbf{V}_1^T$. Applying (5.2.13) yields

$$\begin{aligned} \delta(\mathcal{M},\mathcal{N}) &= \|\mathbf{P}_{\mathcal{M}}(\mathbf{I}-\mathbf{P}_{\mathcal{N}})\|_2 = \left\|\mathbf{U}_1\mathbf{U}_1^T\mathbf{V}_1\mathbf{V}_1^T\right\|_2 = \left\|\mathbf{U}_1\mathbf{U}_1^T\mathbf{V}_1\mathbf{V}_1^T\mathbf{V}_1\right\|_2 \\ &= \left\|\mathbf{U}_1^T\mathbf{V}_1\right\|_2. \end{aligned} \tag{5.13.12}$$

Similarly, $\delta(\mathcal{N},\mathcal{M}) = \left\|\mathbf{U}_2^T\mathbf{V}_2\right\|_2$. The decomposition (5.15.7) remains valid, so

$$\begin{aligned} \|\mathbf{P}_{\mathcal{M}}-\mathbf{P}_{\mathcal{N}}\|_2 &= \max\left\{ \left\|\mathbf{U}_1^T\mathbf{V}_1\right\|_2,\ \left\|\mathbf{U}_2^T\mathbf{V}_2\right\|_2 \right\} \\ &= \max\left\{\delta(\mathcal{M},\mathcal{N}),\ \delta(\mathcal{N},\mathcal{M})\right\} = \operatorname{gap}(\mathcal{M},\mathcal{N}). \end{aligned} \tag{5.13.13}$$

Below is a summary of these and other properties of the gap measure.

Gap Properties

The following statements are true for subspaces $\mathcal{M}, \mathcal{N} \subseteq \mathbb{R}^n$.

- $\operatorname{gap}(\mathcal{M},\mathcal{N}) = \|\mathbf{P}_{\mathcal{M}}-\mathbf{P}_{\mathcal{N}}\|_2$.
- $\operatorname{gap}(\mathcal{M},\mathcal{N}) = \max\left\{ \|(\mathbf{I}-\mathbf{P}_{\mathcal{N}})\mathbf{P}_{\mathcal{M}}\|_2,\ \|(\mathbf{I}-\mathbf{P}_{\mathcal{M}})\mathbf{P}_{\mathcal{N}}\|_2 \right\}$.
- $\operatorname{gap}(\mathcal{M},\mathcal{N}) = 1$ whenever $\dim\mathcal{M} \neq \dim\mathcal{N}$. (5.13.14)
- If $\dim\mathcal{M} = \dim\mathcal{N}$, then $\delta(\mathcal{M},\mathcal{N}) = \delta(\mathcal{N},\mathcal{M})$, and
  - ▷ $\operatorname{gap}(\mathcal{M},\mathcal{N}) = 1$ when $\mathcal{M}^{\perp}\cap\mathcal{N}$ (or $\mathcal{M}\cap\mathcal{N}^{\perp}$) $\neq \mathbf{0}$, (5.13.15)
  - ▷ $\operatorname{gap}(\mathcal{M},\mathcal{N}) < 1$ when $\mathcal{M}^{\perp}\cap\mathcal{N}$ (or $\mathcal{M}\cap\mathcal{N}^{\perp}$) $= \mathbf{0}$. (5.13.16)

Proof of (5.13.14). Suppose that $\dim\mathcal{M} = r$ and $\dim\mathcal{N} = k$, where $r < k$. Notice that this implies that $\mathcal{M}^{\perp}\cap\mathcal{N} \neq \mathbf{0}$, for otherwise the formula for the dimension of a sum (4.4.19) yields

$$n \geq \dim(\mathcal{M}^{\perp}+\mathcal{N}) = \dim\mathcal{M}^{\perp} + \dim\mathcal{N} = n - r + k > n,$$

which is impossible. Thus there exists a nonzero vector $\mathbf{x} \in \mathcal{M}^{\perp}\cap\mathcal{N}$, and by normalization we can take $\|\mathbf{x}\|_2 = 1$. Consequently, $(\mathbf{I}-\mathbf{P}_{\mathcal{M}})\mathbf{x} = \mathbf{x} = \mathbf{P}_{\mathcal{N}}\mathbf{x}$, so $\|(\mathbf{I}-\mathbf{P}_{\mathcal{M}})\mathbf{P}_{\mathcal{N}}\mathbf{x}\|_2 = 1$. This insures that $\|(\mathbf{I}-\mathbf{P}_{\mathcal{M}})\mathbf{P}_{\mathcal{N}}\|_2 = 1$, which implies $\delta(\mathcal{N},\mathcal{M}) = 1$. ■

Proof of (5.13.15). Assume $\dim\mathcal{M} = \dim\mathcal{N} = r$, and use the formula for the dimension of a sum along with $(\mathcal{M}\cap\mathcal{N}^{\perp})^{\perp} = \mathcal{M}^{\perp}+\mathcal{N}$ (Exercise 5.11.5) to conclude that

$$\begin{aligned} \dim\left(\mathcal{M}^{\perp}\cap\mathcal{N}\right) &= \dim\mathcal{M}^{\perp} + \dim\mathcal{N} - \dim\left(\mathcal{M}^{\perp}+\mathcal{N}\right) \\ &= (n-r) + r - \dim\left(\mathcal{M}\cap\mathcal{N}^{\perp}\right)^{\perp} = \dim\left(\mathcal{M}\cap\mathcal{N}^{\perp}\right). \end{aligned}$$

When $\dim\left(\mathcal{M}\cap\mathcal{N}^{\perp}\right) = \dim\left(\mathcal{M}^{\perp}\cap\mathcal{N}\right) > 0$, there are vectors $\mathbf{x}\in\mathcal{M}^{\perp}\cap\mathcal{N}$ and $\mathbf{y}\in\mathcal{M}\cap\mathcal{N}^{\perp}$ such that $\|\mathbf{x}\|_2 = 1 = \|\mathbf{y}\|_2$. Hence, $\|(\mathbf{I}-\mathbf{P}_{\mathcal{M}})\mathbf{P}_{\mathcal{N}}\mathbf{x}\|_2 = \|\mathbf{x}\|_2 = 1$, and $\|(\mathbf{I}-\mathbf{P}_{\mathcal{N}})\mathbf{P}_{\mathcal{M}}\mathbf{y}\|_2 = \|\mathbf{y}\|_2 = 1$, so

$$\delta(\mathcal{N},\mathcal{M}) = \|(\mathbf{I}-\mathbf{P}_{\mathcal{M}})\mathbf{P}_{\mathcal{N}}\|_2 = 1 = \|(\mathbf{I}-\mathbf{P}_{\mathcal{N}})\mathbf{P}_{\mathcal{M}}\|_2 = \delta(\mathcal{M},\mathcal{N}). \quad \blacksquare$$

Proof of (5.13.16). Let $\mathbf{U} = \left(\mathbf{U}_1\,|\,\mathbf{U}_2\right)$ and $\mathbf{V} = \left(\mathbf{V}_1\,|\,\mathbf{V}_2\right)$ be orthogonal matrices defined in (5.13.11), and assume that $\dim\mathcal{M} = \dim\mathcal{N} = r$ with $\dim\left(\mathcal{M}\cap\mathcal{N}^{\perp}\right) = \dim\left(\mathcal{M}^{\perp}\cap\mathcal{N}\right) = 0$. The matrix $\mathbf{U}_2^T\mathbf{V}_1$ is nonsingular because it is $(n-r)\times(n-r)$ and has rank $n-r$ (apply the formula (4.5.1) for the rank of a product). From (5.13.12) we have

$$\begin{aligned}
\delta^2(\mathcal{M},\mathcal{N}) &= \left\|\mathbf{U}_1^T\mathbf{V}_1\right\|_2^2 = \max_{\|\mathbf{x}\|_2=1} \mathbf{x}^T\mathbf{V}_1^T\mathbf{U}_1\mathbf{U}_1^T\mathbf{V}_1\mathbf{x} \\
&= \max_{\|\mathbf{x}\|_2=1} \mathbf{x}^T\mathbf{V}_1^T(\mathbf{I}-\mathbf{U}_2\mathbf{U}_2^T)\mathbf{V}_1\mathbf{x} = \max_{\|\mathbf{x}\|_2=1}\left(1-\left\|\mathbf{U}_2^T\mathbf{V}_1\mathbf{x}\right\|_2^2\right) \\
&= 1 - \min_{\|\mathbf{x}\|_2=1}\left\|\mathbf{U}_2^T\mathbf{V}_1\mathbf{x}\right\|_2^2 = 1 - \frac{1}{\left\|(\mathbf{U}_2^T\mathbf{V}_1)^{-1}\right\|_2^2} < 1 \text{ (recall (5.2.6)).}
\end{aligned}$$

A similar argument shows $\delta^2(\mathcal{N},\mathcal{M}) = \left\|\mathbf{U}_2^T\mathbf{V}_2\right\|_2^2 = 1 - 1/\left\|(\mathbf{U}_2^T\mathbf{V}_1)^{-1}\right\|_2^2$ (Exercise 5.15.11(b)), so $\delta(\mathcal{N},\mathcal{M}) = \delta(\mathcal{M},\mathcal{N}) < 1$. $\blacksquare$

Because $0 \le \text{gap}\,(\mathcal{M},\mathcal{N}) \le 1$, the gap measure defines another angle between $\mathcal{M}$ and $\mathcal{N}$.

Maximal Angle

The ***maximal angle*** between subspaces $\mathcal{M},\,\mathcal{N}\subseteq\mathbb{R}^n$ is defined to be the number $0\le\theta_{max}\le\pi/2$ for which

$$\sin\theta_{max} = \text{gap}\,(\mathcal{M},\mathcal{N}) = \|\mathbf{P}_{\mathcal{M}}-\mathbf{P}_{\mathcal{N}}\|_2\,. \tag{5.13.17}$$

For applications requiring knowledge of the degree of separation between a pair of nontrivial complementary subspaces, the minimal angle does the job. Similarly, the maximal angle adequately handles the task for subspaces of equal dimension. However, neither the minimal nor maximal angle may be of much help for more general subspaces. For example, if $\mathcal{M}$ and $\mathcal{N}$ are subspaces of unequal dimension that have a nontrivial intersection, then $\theta_{min} = 0$ and $\theta_{max} = \pi/2$, but neither of these numbers might convey the desired information. Consequently, it seems natural to try to formulate definitions of "intermediate" angles between θ_{min} and θ_{max}. There are a host of such angles known as the ***principal*** or ***canonical angles***, and they are derived as follows.