5.6 Differentiate f with respect to t and g with respect to $\mathbf{X}$, where

$$f(\mathbf{t}) = sin(log(\mathbf{t}^T\mathbf{t})), \qquad \mathbf{t} \in \mathbb{R}^D$$

$$g(\mathbf{X}) = tr(\mathbf{AXB}), \qquad \mathbf{A} \in \mathbb{R}^{D\times E}, \mathbf{X} \in \mathbb{R}^{E\times F}, \mathbf{B} \in \mathbb{R}^{F\times D}$$

Solution.

$$f'(\mathbf{t}) = \frac{d(sin(log(\mathbf{t}^T\mathbf{t})))}{d\mathbf{t}}$$

$$= \frac{d(sin(log(\mathbf{t}^T\mathbf{t})))}{d(log(\mathbf{t}^T\mathbf{t}))} * \frac{d(log(\mathbf{t}^T\mathbf{t}))}{d(\mathbf{t}^T\mathbf{t})} * \frac{d(\mathbf{t}^T\mathbf{t})}{d\mathbf{t}}$$

$$= cos(log(\mathbf{t}^T\mathbf{t}) * (\frac{1}{\mathbf{t}^T\mathbf{t}}) * 2\mathbf{t}^T$$

$$g'(\mathbf{X}) = \frac{d(tr(\mathbf{AXB}))}{d\mathbf{X}} = tr\left(\frac{d(\mathbf{AXB})}{d\mathbf{X}}\right) \qquad \text{Using property 5.100}$$

$$\text{Now, } \mathbf{AXB} = \begin{bmatrix} \mathbf{a}_1 \\ \mathbf{a}_2 \\ \vdots \\ \mathbf{a}_D \end{bmatrix} \begin{bmatrix} & X & \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_D \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{a}_1 X \\ \mathbf{a}_2 X \\ \vdots \\ \mathbf{a}_D X \end{bmatrix} \begin{bmatrix} \mathbf{b}_1 & \mathbf{b}_2 & \dots & \mathbf{b}_D \end{bmatrix}$$

$$= \begin{bmatrix} \mathbf{a}_1 X \mathbf{b}_1 & \mathbf{a}_1 X \mathbf{b}_2 & \dots & \mathbf{a}_1 X \mathbf{b}_D \\ \mathbf{a}_2 X \mathbf{b}_1 & \mathbf{a}_2 X \mathbf{b}_2 & \dots & \mathbf{a}_2 X \mathbf{b}_D \\ \vdots & \vdots & \dots & \vdots \\ \mathbf{a}_D X \mathbf{b}_1 & \mathbf{a}_D X \mathbf{b}_2 & \dots & \mathbf{a}_D X \mathbf{b}_D \end{bmatrix}$$

Now, $tr(\mathbf{AXB}) = \Sigma_{i=1}^{D} \mathbf{a}_i \mathbf{X} \mathbf{b}_i$

$$\implies \frac{d(tr(\mathbf{AXB}))}{d\mathbf{X}} = \frac{d(\Sigma_{i=1}^{D} \mathbf{a}_i \mathbf{X} \mathbf{b}_i)}{d\mathbf{X}} = \Sigma_{i=1}^{D} \mathbf{a}_i^T \mathbf{b}_i^T \qquad \text{Using property 5.106}$$

Since $\mathbf{a}_i$ was a row vector and $\mathbf{b}_i$ a column vector, $\mathbf{a}_i^T \mathbf{b}_i^T$ is an outer product.

The multiplication of 2 matrices can be represented as the sum of the the outer products of their constituent vectors. Source :

`http://mlwiki.org/index.php/Matrix-Matrix_Multiplication`

So, we get $\Sigma_{i=1}^{D} \mathbf{a}_i^T \mathbf{b}_i^T = \mathbf{A}^T \mathbf{B}^T \implies g'(\mathbf{X}) = \mathbf{A}^T \mathbf{B}^T$