Optimality of Kurtosis Contrast Function in ICA Optimization with Orthogonal Constraint

For whitened data, *e.g.*, $E[\mathrm{x}\mathrm{x}^{\mathrm{T}}] = \mathrm{I}$, the kurtosis contrast function is

$$
\begin{aligned}
f(\mathrm{w}) = kurt(\mathrm{w}^{\mathrm{T}}\mathrm{x}) &= \frac{E[(\mathrm{w}^{\mathrm{T}}\mathrm{x})^4]}{E^2[(\mathrm{w}^{\mathrm{T}}\mathrm{x})^2]} - 3 \\
&= \frac{E[(\mathrm{w}^{\mathrm{T}}\mathrm{x})^4]}{E^2[\mathrm{w}^{\mathrm{T}}\mathrm{x}\mathrm{x}^{\mathrm{T}}\mathrm{w}\ ]} - 3 \\
&= \frac{E[(\mathrm{w}^{\mathrm{T}}\mathrm{x})^4]}{(\mathrm{w}^{\mathrm{T}}E[\mathrm{x}\mathrm{x}^{\mathrm{T}}]\mathrm{w})^2} - 3 \\
&= \frac{E[(\mathrm{w}^{\mathrm{T}}\mathrm{x})^4]}{\|\mathrm{w}\|^4} - 3
\end{aligned}
$$

Take the derivative *w.r.t.* w to get the gradient vector,

$$
\begin{aligned}
\nabla f(\mathrm{w}) &= \frac{\nabla E[(\mathrm{w}^{\mathrm{T}}\mathrm{x})^4]\,\|\mathrm{w}\|^4 - E[(\mathrm{w}^{\mathrm{T}}\mathrm{x})^4]\cdot 4\cdot \|\mathrm{w}\|^2\cdot \mathrm{w}}{(\|\mathrm{w}\|^4)^2} \\
&= \frac{4E[(\mathrm{w}^{\mathrm{T}}\mathrm{x})^3\mathrm{x}]}{\|\mathrm{w}\|^4} - \frac{4E[(\mathrm{w}^{\mathrm{T}}\mathrm{x})^4]}{\|\mathrm{w}\|^6}\mathrm{w}
\end{aligned}
$$

where $\nabla E[(\mathrm{w}^{\mathrm{T}}\mathrm{x})^4] = E[4\cdot(\mathrm{w}^{\mathrm{T}}\mathrm{x})^3\cdot \mathrm{x}]$.

Calculate the inner product of the demixing vector and the gradient vector:

$$
\begin{aligned}
\mathrm{w}^{\mathrm{T}}\nabla f(\mathrm{w}) &= \frac{4E[(\mathrm{w}^{\mathrm{T}}\mathrm{x})^3\mathrm{w}^{\mathrm{T}}\mathrm{x}]}{\|\mathrm{w}\|^4} - \frac{4E[(\mathrm{w}^{\mathrm{T}}\mathrm{x})^4]\mathrm{w}^{\mathrm{T}}\mathrm{w}}{\|\mathrm{w}\|^6} \\
&= \frac{4E[(\mathrm{w}^{\mathrm{T}}\mathrm{x})^4]}{\|\mathrm{w}\|^4} - \frac{4E[(\mathrm{w}^{\mathrm{T}}\mathrm{x})^4]}{\|\mathrm{w}\|^4} \\
&= 0
\end{aligned}
$$

Therefore, we have:

$\mathrm{w} \perp \nabla f(\mathrm{w})$. □

Yiou Li, ENEE718 Final Project Reference
4/15/2005