PCA dual problem (cont.)

Solution: Assume $D^* = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_n \end{pmatrix}$ is of full rank,

i.e. $\lambda_1 \geq \cdots \geq \lambda_n > 0$

Then let $\breve{B}_{d \times n} = \tilde{X} B^* \left(D^*\right)^{-1/2}$,

Where

$$\left(D^*\right)^{-1/2} = \begin{pmatrix} \lambda_1^{1/2} & & 0 \\ & \ddots & \\ 0 & & \lambda_n^{1/2} \end{pmatrix}$$

PCA dual problem (cont.)

And “fill out the rest of B” with “columns in null space”,

I.e. let $G_{d\times(d-n)}$ be $d-n$ orthonormal column vectors,

that are orthogonal to $\tilde{X}$ (compute by Gram-Schmidt process)

Thus “pad $\breve{B}$ out to a basis matrix”, by defining:

$$B = \begin{pmatrix} \breve{B} & G \end{pmatrix}$$

PCA dual problem (cont.)

Check orthonormality:

$$B^t B = \begin{pmatrix} \breve{B}^t \\ G^t \end{pmatrix} \begin{pmatrix} \breve{B} & G \end{pmatrix} = \begin{pmatrix} \breve{B}^t \breve{B} & 0 \\ 0 & I \end{pmatrix}$$

but

$$\breve{B}^t \breve{B} = \left(\tilde{X} B^* \left(D^*\right)^{-1/2} \right)^t \left(\tilde{X} B^* \left(D^*\right)^{-1/2} \right) = \left(\left(D^*\right)^{-1/2} B^{*t} \tilde{X}^t \right) \left(\tilde{X} B^* \left(D^*\right)^{-1/2} \right)$$
$$\breve{B}^t \breve{B} = \left(D^*\right)^{-1/2} B^{*t} \left(\tilde{X}^t \tilde{X} \right) B^* \left(D^*\right)^{-1/2} = \left(\left(D^*\right)^{-1/2} B^{*t} \right) \Sigma^* \left(B^* \left(D^*\right)^{-1/2} \right)$$
$$\breve{B}^t \breve{B} = \left(D^*\right)^{-1/2} \left(B^{*t} \Sigma^* B^* \right) \left(D^*\right)^{-1/2} = \left(D^*\right)^{-1/2} D^* \left(D^*\right)^{-1/2} = I$$

so B is orthonormal.

PCA dual problem (cont.)

Check diagonalization:

$$B^t\hat{\Sigma}B = \begin{pmatrix} \breve{B}^t \\ G^t \end{pmatrix}\hat{\Sigma}\begin{pmatrix} \breve{B} & G \end{pmatrix} = \begin{pmatrix} \breve{B}^t \\ G^t \end{pmatrix}\begin{pmatrix} \hat{\Sigma}\breve{B} & \hat{\Sigma}G \end{pmatrix} = \begin{pmatrix} \breve{B}^t\hat{\Sigma}\breve{B} & \breve{B}^t\hat{\Sigma}G \\ G^t\hat{\Sigma}\breve{B} & G^t\hat{\Sigma}G \end{pmatrix}$$

but

$$\begin{aligned}
\breve{B}^t\hat{\Sigma}\breve{B} &= \left(D^*\right)^{-1/2} B^{*t}\,\tilde{X}^t\hat{\Sigma}\tilde{X}B^*\left(D^*\right)^{-1/2} = \\
&= \left(D^*\right)^{-1/2} B^{*t}\,\tilde{X}^t\,\tilde{X}\tilde{X}^t\,\tilde{X}B^*\left(D^*\right)^{-1/2} = \\
&= \left(D^*\right)^{-1/2} B^{*t}\Sigma^*\Sigma^*B^*\left(D^*\right)^{-1/2} = \left(D^*\right)^{-1/2} B^{*t}\Sigma^*\left(B^*B^{*t}\right)\Sigma^*B^*\left(D^*\right)^{-1/2} = \\
&= \left(D^*\right)^{-1/2}\left(B^{*t}\Sigma^*B^*\right)\left(B^{*t}\Sigma^*B^*\right)\left(D^*\right)^{-1/2} = \left(D^*\right)^{-1/2}D^*D^*\left(D^*\right)^{-1/2} = D^*
\end{aligned}$$

[go to next page]