Derivative of SVD for complex-valued matrices

$A \in \mathbb{C}^{m\times n}$ of rank $k \leq \min(m,n) \implies A = USV^\dagger$ with $U \in \mathbb{C}^{m\times k}$,

$S \in \mathbb{R}^{k\times k}$ diagonal, $V \in \mathbb{C}^{n\times k}$; and $U^\dagger U = V^\dagger V = \mathbb{1}_k$ (i) (semi-unitary)

$\Rightarrow dA = dU\, S V^\dagger + U\, dS\, V^\dagger + U S\, dV^\dagger$ (ii)

Differentiating constraint (i) yields $dU^\dagger U + U^\dagger dU = 0$; $dV^\dagger V + V^\dagger dV = 0$

$\implies d\Omega_U := U^\dagger dU$ and $d\Omega_V := V^\dagger dV$ are $k\times k$ anti-hermitean matrices!

We can always find matrices $U_\perp \in \mathbb{C}^{m\times(m-k)}$ and $V_\perp \in \mathbb{C}^{n\times(n-k)}$ such that $[U\ U_\perp]$ and $[V\ V_\perp]$ are unitary (unitary completion, e.g. via Gram-Schmidt).

We may then expand dU as $dU = U\, d\Omega_U + U_\perp\, dK_U$ with $dK_U \in \mathbb{C}^{(m-k)\times k}$ unconstrained.
and expand dV as $dV = V\, d\Omega_V + V_\perp\, dK_V$ with $dK_V \in \mathbb{C}^{(n-k)\times k}$ unconstrained

We proceed by left-multiplying (ii) by $U^\dagger$ and right-multiplying by V:

$$dP := U^\dagger dA\, V = U^\dagger dU\, S + dS + S\, dV^\dagger V$$
$$= d\Omega_U S + dS + S\, d\Omega_V^\dagger$$

$dU = U\, d\Omega_U + U_\perp\, dK_U$
$dV^\dagger = d\Omega_V^\dagger V^\dagger + dK_V^\dagger V_\perp^\dagger$

Since $d\Omega_U$ and $d\Omega_V$ are anti-hermitean, their diagonal elements are <u>purely imaginary</u>. However, dS is a diagonal matrix with <u>purely real entries</u>. Thus its clear that

because singular values are always real!

$$dS = \mathbb{1}_k \circ \mathrm{Re}[dP]$$

Hadamard product

and $d\Omega_U S - S\, d\Omega_V = \mathbb{1}_k \circ i\, \mathrm{Im}[dP] + \bar{\mathbb{1}}_k \circ dP$ (iii)

Taking the complex conjugate transpose yields $-S\, d\Omega_U + d\Omega_V S = -\mathbb{1}_k \circ i\, \mathrm{Im}[dP] + \bar{\mathbb{1}}_k \circ dP^\dagger$ (iv)

To proceed, we right-multiply (iii) by S and left-multiply (iv) by S and add them together, yielding

$$d\Omega_u S^2 - \cancel{S d\Omega_v S} - S^2 d\Omega_u + \cancel{S d\Omega_v S} = \left[\mathbb{1}_k \circ i\,\mathrm{Im}[dP]\right] S - S\left(\mathbb{1}_k \circ i\,\mathrm{Im}[dP]\right) + \left[\bar{\mathbb{1}}_k \circ dP\right] S + S\left[\bar{\mathbb{1}}_k \circ dP^\dagger\right]$$

$$\Longrightarrow d\Omega_u S^2 - S^2 d\Omega_v = \bar{\mathbb{1}}_k \circ S dP + \bar{\mathbb{1}}_k \circ dP^\dagger S$$

$$D(A \circ B) = DA \circ B = A \circ DB$$
$$(A \circ B) D = AD \circ B = A \circ BD$$

$$= \bar{\mathbb{1}}_k \circ \left[dP S + S dP^\dagger\right]$$

$$A \circ B + A \circ C = A \circ (B + C)$$

$$\Longrightarrow \bar{\mathbb{1}}_k \circ \left[dP S + S dP^\dagger\right] = d\Omega_u S^2 - S^2 d\Omega_v$$

$$\bar{\mathbb{1}}_k \circ Z \overset{!}{=} XD - DX$$ with $Z^\dagger = Z$, $X^\dagger = -X$, $D \in \mathbb{R}^{k \times k}$ diagonal, solve for X.

$X = F \circ Z$ with $F_{ij} := \begin{cases} \frac{1}{d_j - d_i} & i \neq j \\ 0 & i = j \end{cases}$ $\Longrightarrow$ $X^\dagger = F^\dagger \circ Z^\dagger = -F \circ Z$ is skew-symmetric. ✓

$$XD - DX = (F \circ Z)D - D(F \circ Z) = (FD) \circ Z - (DF) \circ Z = (*)$$

Let $i \neq j$. then
$$(FD)_{ij} = \sum_\mu F_{i\mu} d_{\mu j} = \sum_\mu \frac{1}{d_\mu - d_i} \delta_{\mu j} d_j = \frac{d_j}{d_j - d_i}$$
$$(DF)_{ij} = \sum_\mu d_{i\mu} F_{\mu j} = \sum_\mu d_i \delta_{i\mu} \frac{1}{d_j - d_\mu} = \frac{d_i}{d_j - d_i}$$

also, $(FD)_{ii} = (DF)_{ii} = 0$

$$\Longrightarrow (FD)_{ij} = \begin{cases} \frac{d_j}{d_j - d_i} & i \neq j \\ 0 & i = j \end{cases}, \quad (DF)_{ij} = \begin{cases} \frac{d_i}{d_j - d_i} & i \neq j \\ 0 & i = j \end{cases}$$

$$(*) = (FD - DF) \circ Z = \bar{\mathbb{1}}_k \circ Z \quad ✓$$

Therefore, the equation is solved by $\boxed{d\Omega_u = F \circ \left[dP S + S dP^\dagger\right]}$ (v)

We can obtain a similar equation for $d\Omega_v$ by left-multiplying (iii) by S and right-multiplying (iv) by S before adding:

$$\left.\begin{aligned} S d\Omega_u S - S^2 d\Omega_v &= S \circ i\,\mathrm{Im}[dP] + \bar{\mathbb{1}}_k \circ S dP \\ -S d\Omega_u S + d\Omega_v S^2 &= -S \circ i\,\mathrm{Im}[dP] + \bar{\mathbb{1}}_k \circ dP^\dagger S \end{aligned}\right\} \Longrightarrow d\Omega_v S^2 - S^2 d\Omega_v = \bar{\mathbb{1}}_k \circ \left[S dP + dP^\dagger S\right]$$

solved by $\boxed{d\Omega_v = F \circ \left[S dP + dP^\dagger S\right]}$ (vi)

Equations (v) and (vi) do not fix the diagonals of $d\Omega_U$ and $d\Omega_V$, which are purely imaginary.
We can fix this by looking at the diagonal part of eq. (iii):

$$\mathbb{1}_k \circ \left[d\Omega_U S - S d\Omega_V \right] = \mathbb{1}_k \circ i\,\mathrm{Im}\left[dP\right]$$

$$\Rightarrow \quad \mathbb{1}_k \circ \left[d\Omega_U - d\Omega_V \right] = \mathbb{1}_k \circ i\,\mathrm{Im}\left[dP\right] S^{-1}$$

We can satisfy this equation by setting

I'm choosing this solution because of symmetry. I don't know if it actually matters how we „distribute" the dD terms between $d\Omega_U$ and $d\Omega_V$

$$\boxed{\begin{aligned} d\Omega_U &= F \circ \left[dP S + S dP^\dagger\right] + \mathbb{1}_k \circ dD \\ d\Omega_V &= F \circ \left[S dP + dP^\dagger S\right] - \mathbb{1}_k \circ dD \end{aligned} \quad \text{with } dD := \frac{i\,\mathrm{Im}[dP]\,S^{-1}}{2}}$$

(Vii)

We are left with finding dK_U and dK_V. Multiplying (ii) by $U_\perp^\dagger$ yields

$$U_\perp^\dagger dA = U_\perp^\dagger dU S V^\dagger = U_\perp^\dagger \left(U d\Omega_U + U_\perp dK_U \right) S V^\dagger = dK_U S V^\dagger \qquad | \cdot V \cdot S^{-1}$$

$$\boxed{dK_U = U_\perp^\dagger dA\, V S^{-1}}$$

(Vii)

Similarly, we can left-multiply $(ii)^\dagger$ by $V_\perp^\dagger$ to obtain

$$V_\perp^\dagger dA^\dagger = V_\perp^\dagger dV S U^\dagger = V_\perp^\dagger \left(V d\Omega_V + V_\perp dK_V \right) S U^\dagger = dK_V S U^\dagger \quad | \cdot U \cdot S^{-1}$$

$$\boxed{dK_V = V_\perp^\dagger dA^\dagger U S^{-1}}$$

(Viii)

Using all previous derivations, we can finally write

$$\begin{aligned} dU &= U \left(F \circ \left[dP S + S dP^\dagger \right] + \mathbb{1} \circ dD \right) + U_\perp U_\perp^\dagger dA\, V S^{-1} \\ dS &= \mathbb{1}_k \circ \mathrm{Re}\left[dP\right] \\ dV &= V \left(F \circ \left[S dP + dP^\dagger S \right] - \mathbb{1} \circ dD \right) + V_\perp V_\perp^\dagger dA^\dagger U S^{-1} \end{aligned}$$

Using the fact that $M_U = [U\ U_\perp]$ and $M_V = [V\ V_\perp]$ are unitary:

$$\Rightarrow M_U M_U^\dagger = [U\ U_\perp]\begin{bmatrix} U^\dagger \\ U_\perp^\dagger \end{bmatrix} = UU^\dagger + U_\perp U_\perp^\dagger = \mathbb{1} \ ; \ M_V M_V^\dagger = [V\ V_\perp]\begin{bmatrix} V^\dagger \\ V_\perp^\dagger \end{bmatrix} = VV^\dagger + V_\perp V_\perp^\dagger = \mathbb{1}$$

and we can rewrite

$$\boxed{\begin{aligned} dU &= U\left(F \circ [dP S + S dP^\dagger] + \mathbb{1}_k \circ dD\right) + \left(\mathbb{1}_m - UU^\dagger\right) dA\, V S^{-1} \\ dS &= \mathbb{1}_k \circ \mathrm{Re}[dP] \\ dV &= V\left(F \circ [S dP + dP^\dagger S] - \mathbb{1}_k \circ dD\right) + \left(\mathbb{1}_n - VV^\dagger\right) dA^\dagger\, U S^{-1} \end{aligned}}$$

There is one small subtlety left to discuss. The SVD is not unique, but has a gauge degree of freedom: $A = USV^\dagger = U'SV'^\dagger = (U\Lambda)S(V\Lambda)^\dagger = U\Lambda S\Lambda^\dagger V = US(\Lambda\Lambda^\dagger)V$

with $\Lambda_{jj} = e^{i\theta_j}$ and $\Lambda_{ij} = 0$ for $i \neq j$. $\Rightarrow$ k degrees of freedom!

$\Rightarrow$ „dU" is illdefined, we can only really define derivatives of functions that are gauge invariant!