Optimised photometric stereo via non-convex variational minimisation (Supplementary Material)

Laurent Hoeltgen[1]
hoeltgen@b-tu.de
Yvain Quéau[2]
queau@enseeiht.fr
Michael Breuß[1]
breuss@b-tu.de
Georg Radow[1]
radow@b-tu.de

[1] Chair for Applied Mathematics
BTU Cottbus-Senftenberg,
Cottbus, Germany

[2] IRIT, UMR CNRS 5505
Université de Toulouse
Toulouse, France

In the following paragraphs we present the proofs that were omitted in the main paper due to spatial constraints. The results found in this document complement the statements given in Section 3.6 of our paper.

The following two lemmas state extensions of the product and chain-rule to matrix valued settings. They provide us closed form representations that will be useful for the forthcoming findings. These results have been extracted from [2] (Theorem 7 and 9 respectively). Since these lemmas have been copied verbatim, we refer to their source for the detailed proofs.

Lemma 1 (Chain Rule for Matrix Differentials)**.** *Let S be a subset of $\mathbb{R}^{n,q}$ and assume that $F\colon S \to \mathbb{R}^{m,p}$ is differentiable at an interior point C of S. Let T be a subset of $\mathbb{R}^{m,p}$ such that $F(X) \in T$ for all $X \in S$, and assume that $G\colon T \to \mathbb{R}^{r,s}$ is differentiable at an interior point $B = F(C)$ of T. Then the composite function $H\colon S \to \mathbb{R}^{r,s}$ defined by $H(X) = G(F(X))$ is differentiable at C and*

$$D[H](C) = D[G](B)D[F](C) \tag{1}$$

Lemma 2 (Product Rule for Matrix Differentials)**.** *Let $U\colon S \to \mathbb{R}^{m,r}$ and $V\colon S \to \mathbb{R}^{r,p}$ be two matrix functions defined and differentiable on an open set $S \subseteq \mathbb{R}^{n,q}$. Then the matrix product $UV\colon S \to \mathbb{R}^{m,p}$ is differentiable on S and the Jacobian matrix $D[UV](X) \in \mathbb{R}^{mp,nq}$ is given by*

$$D\left[UV\right](X) = (V^{\top} \otimes I_m)D[U](X) + (I_p \otimes U)D[V](X) \tag{2}$$

Here, I_k represents the identity matrix in $\mathbb{R}^{k,k}$ and the symbol $\otimes$ stands for the Kronecker matrix product, e.g. [1] (Definition 4.2.1).

The following two corollaries are a direct consequence from the foregoing statements. It suffices to plug in the corresponding quantities. We also remind, that our choice of the matrix derivative allows us to interpret vectors as matrices having a single column only.

Corollary 1. *Let $A(z)$ be an $n \times q$ matrix depending on $z \in \mathbb{R}^m$ and $M \in \mathbb{R}^{q,m}$ a matrix which does not depend on z, then the Jacobian of the matrix-vector product $A(z)Mz$ is given by*

$$D[A(z)Mz](z) = \left((Mz)^\top \otimes I_n\right) D[A](z) + A(z)M \tag{3}$$

Proof. We apply the product rule on the product between $A(z)$ and Mz and subsequently on the product Mz. In a first step this yields

$$D[A(z)Mz] = \left((Mz)^\top \otimes I_n\right) D[A](z) + A(z)D[Mz](z) \tag{4}$$

Since $D[Mz](z) = M$ the result follows immediately. □

Corollary 2 and Theorem 1 yield our desired compact representations that we used in our main paper for the algorithmic presentation of our iterative schemes.

Corollary 2. *Using the same assumptions as in Corollary 1, we deduce from the chain rule given in Lemma 1 the following relationship*

$$\nabla \|A(z)Mz\|_2^2 = 2\big(\underbrace{D[A](z)^\top\left((Mz) \otimes I_n\right) + M^\top A(z)^\top}_{=D[A(z)Mz](z)^\top}\big)A(z)Mz \tag{5}$$

Proof. Since $D\left[\|x\|_2^2\right](x)$ is given by $2x^\top$ we conclude from the chain- and product-rule that

$$\begin{aligned} D[\|A(z)Mz\|_2^2](z) &= 2\big(A(z)Mz\big)^\top D[A(z)Mz](z) \\ &= 2\big(A(z)Mz\big)^\top \big(((Mz)^\top \otimes I_n)D[A](z) + A(z)D[Mz](z)\big) \end{aligned} \tag{6}$$

Since the gradient is simply the transposed version of the Jacobian, we obtain

$$\nabla \|A(z)Mz\|_2^2 = 2\left(\left((Mz)^\top \otimes I_n\right) D[A](z) + A(z)D[Mz](z)\right)^\top A(z)Mz \tag{7}$$

from which the statement follows immediately. □

Let us now come to our main result. It proves the claim in Theorem 1 from our paper. We remark, that the formulation in our paper has an additional multiplicative factor $\frac{1}{2m}$, which we omitted in this presentation.

Theorem 1. *Let $f(z) = \|A(z)Mz - b(z)\|_2^2$ be given with sufficiently smooth data $A(z)$ and $b(z)$. Then we have for the gradient of f the following closed form expression:*

$$\nabla f(z) = 2\left(A(z)M + \left((Mz)^\top \otimes I_n\right) D[A](z) - D[b](z)\right)^\top (A(z)Mz - b(z)) \tag{8}$$

Proof. From the relationship between the canonical scalar product in $\mathbb{R}^n$ and the Euclidean norm we deduce that

$$\|A(z)Mz - b(z)\|_2^2 = \|A(z)Mz\|_2^2 + \|b(z)\|_2^2 - 2\langle A(z)Mz, b(z)\rangle \tag{9}$$

Applying the gradient at each term separately and using the results from the previous corollaries, we obtain

$$\begin{aligned}\nabla \|A(z)Mz - b(z)\|_2^2 = \\ 2\underbrace{\left(D[A](z)^\top\left((Mz)\otimes I_n\right)+M^\top A(z)^\top\right)}_{=D[A(z)Mz](z)^\top}A(z)Mz + 2D[b](z)^\top b(z) - \\ 2\underbrace{\left(D[A](z)^\top\left((Mz)\otimes I_n\right)+M^\top A(z)^\top\right)}_{=D[A(z)Mz](z)^\top}b(z) - 2D[b](z)^\top A(z)Mz\end{aligned} \tag{10}$$

which can be simplified to

$$\nabla \|A(z)Mz - b(z)\|_2^2 = 2\left(D[A(z)Mz](z) - D[b](z)\right)^\top \left(A(z)Mz - b(z)\right) \tag{11}$$

The result follows now from the linearity of the Jacobian. □

References

[1] R. A. Horn and C. R. Johnson. *Topics in Matrix Analysis*. Cambridge University Press, 1994.

[2] J. R. Magnus and H. Neudecker. Matrix differential calculus with applications to simple, Hadamard, and Kronecker products. *Journal of Mathematical Psychology*, 29:474–492, 1985.