Conventional solution and proof

The source's "Example 18.12" and "Exercise 18.13" are references to the numbered equations. Using their exact SVD gives

$$C_1 = \sqrt{3} u_1 v_1^\top = \begin{pmatrix} 1 & 1 \\ 1/2 & 1/2 \\ 1/2 & 1/2 \end{pmatrix}.$$

The residual is $u_2 v_2^\top$, so its Frobenius norm is $\|u_2\|\|v_2\| = 1$. Directly, its four nonzero entries have magnitude $1/2$, and their squared sum is one. To prove optimality here without merely citing it, let B have rank at most one. Its restriction to the two-dimensional span of v_1, v_2 has a nonzero kernel vector; choose a unit vector w in that kernel. Then $\|(C - B)w\| = \|Cw\| \geq \sigma_2 = 1$, because $\|Cw\|^2 = 3|\langle w, v_1\rangle|^2 + |\langle w, v_2\rangle|^2 \geq 1$. Since $\|C - B\|_F \geq \|C - B\|_2 \geq 1$, the displayed approximation attains the least possible error.

Structural or analytic derivation

Orthogonal singular components isolate the discarded energy exactly. The lower-bound witness proves optimality over every rank-one competitor, not just over approximations obtained by deleting the second term.

Comparative cost and preferred method

A rank-one factorization stores $M + N + 1$ scalars and applies to a vector in $O(M + N)$ arithmetic. For this 3×2 example the factor representation is not smaller than the six original entries; it illustrates the theorem, not a practical compression win at this size.