construction given a linear map $A : \mathbf{R}^n \to \mathbf{R}^m$, we construct a function $A^\dagger : \mathbf{R}^m \to \mathbf{R}^n$ as follows. given $b \in \mathbf{R}^m$ let $A^\dagger(b)$ be the vector $x \in \mathbf{R}^n$ of minimal norm amongst all vectors minimizing $\|Ax - b\|$.

observations
for A is invertible we have $A^\dagger = A^{-1}$
for A an orthogonal projection we have $A^\dagger = A$

exercise if $x^* \in \mathbf{R}^n$ is any minimizer of $\|Ax - b\|$ then $A^\dagger(b) = \mathrm{Proj}_{(\ker A)^\perp} x^*$. in particular, in order to find $A^\dagger(b)$ one must first find $p = \mathrm{Proj}_{\mathrm{img} A} b$, then take any preimage $Ax^* = p$, and finally get $A^\dagger(b) = \mathrm{Proj}_{(\ker A)^\perp} x^*$.

corollary $A^\dagger$ is a linear map. in other words $\cdot^\dagger$ is a function taking $m \times n$ matrices to $n \times m$ matrices.

observation we have $AA^\dagger = \mathrm{Proj}_{\mathrm{img} A}$ and $A^\dagger A = \mathrm{Proj}_{(\ker A)^\perp}$.
in particular, if the columns of A are linearly independent then $A^\dagger A = I$ and if the rows of A are linearly independent then $AA^\dagger = I$.

corollary we have
i. $AA^\dagger A = A$
ii. $A^\dagger A A^\dagger = A^\dagger$
iii. $AA^\dagger$ and $A^\dagger A$ are symmetric

claim $A^\dagger$ is characterized by the above three properties.

proof suppose $B \in \mathbf{R}^{n \times m}$ also satisfies these properties. we get $AB = AA^\dagger AB = (AA^\dagger)^t (AB)^t = (A^\dagger)^t (ABA)^t = AA^\dagger$ and similarly $BA = A^\dagger A$. finally $B = BAB = A^\dagger AB = A^\dagger AA^\dagger = A^\dagger$.

corollaries
$\cdot^\dagger$ is an involution, namely $(A^\dagger)^\dagger = A$
$(\ker A)^\perp = \mathrm{img} A^t = \mathrm{img} A^\dagger$
$(A^\dagger)^t = (A^t)^\dagger$
if $A = \mathrm{diag}(d_i)$ then $A^\dagger = \mathrm{diag}(b_i)$ where $b_i = \begin{cases} 1/d_i & d_i \neq 0 \\ 0 & d_i = 0 \end{cases}$
if U is orthogonal then $(UA)^\dagger = A^\dagger U^t$
if V is orthogonal then $(AV)^\dagger = V^t A^\dagger$

conclusion
$x = A^\dagger b$ is a minimizer for $\|Ax - b\|$. we may easily compute $A^\dagger$ via the SVD decomposition of A.