Book: Bifurcation Analysis of Fluid Flows
Authors: Fred W. Wubs and Henk A. Dijkstra
Chapter: 3, Exercise: 3.10
Exercise author: G. Tiesinga
Version: 1

a. 1. $J_1(x) = \|x-z\|_A = \sqrt{(x-z, A(x-z))}$ so we minimize $(x-z, A(x-z))$.
Using the directional derivative in the direction of arbitrary vector $y \in \mathcal{V}$, we obtain

$$\frac{d}{d\epsilon}(x-z+\epsilon y, A(x-z+\epsilon y))\Big|_{\epsilon=0} = 2(y, A(x-z)) + 2\epsilon(y, Ay)\Big|_{\epsilon=0} = 2(y, A(x-z)) = 0$$

(note: for this inner product, $(x,y) = (y,x)$ and $(x, Ay) = (A^T x, y)$, from the latter we see that if A symmetric we have $(x, Ay) = (Ax, y)$)

Using $\nabla J_1(x) = 0$ approach:

$$\nabla J_1(x) = \nabla(x-z, A(x-z)) = \nabla(x-z)^T A(x-z) = \ldots = 2A(x-z) \ = 0$$

(this uses $\nabla c^T x = c, \nabla x^T A x = (A + A^T)x = 2Ax$ the latter equality only if B symmetric)

Hence, from both approaches we obtain $A(x-z) = 0$ or $Ax = Az$,
or in weak form: find a $x_{\mathcal{V}} \in \mathcal{V}$ s.t. $(A(z - x_{\mathcal{V}}), y) = (z - x_{\mathcal{V}}, Ay) = 0 \ \ \forall y \in \mathcal{V}$.

2. minimize $J_2(x) = \|b - Ax\|^2 = (b - Ax, b - Ax)$

Directional derivative approach: for y arbitrary, show

$$\tfrac{d}{d\epsilon} J_2(x+\epsilon y)\Big|_{\epsilon=0} = \tfrac{d}{d\epsilon}(b - A(x+\epsilon y), b - A(x+\epsilon y))\Big|_{\epsilon=0} = \ldots = -2(Ay, b - Ax) = 0$$

Using $\nabla J_2(x) = 0$ approach: show

$$\nabla J_2(x) = \nabla(b - Ax, b - Ax) = \nabla(b - Ax)^T(b - Ax) = \ldots = 2(A^T Ax, A^T b) = 0$$

Hence, from both approaches we obtain $A^T Ax = A^T b$,
or in weak form: find a $x_{\mathcal{V}} \in \mathcal{V}$ s.t. $(Ay, Ax_{\mathcal{V}} - b) = 0 \ \ \forall y \in \mathcal{V}$.

3. minimize $J_3(x) = \frac{1}{2}(x, Ax) - (b, x)$

Directional derivative approach: for y arbitrary, show

$$\frac{d}{d\epsilon}\frac{1}{2}(x+\epsilon y, A(x+\epsilon y)) - (b, x+\epsilon y)|_{\epsilon=0} = (x, Ay) + \epsilon(y, Ay) - (b, y)|_{\epsilon=0}$$
$$= (x, Ay) - (b, y) = (y, Ax - b) = 0$$

Using $\nabla J_3(x) = 0$ approach: show $\nabla J_3(x) = \nabla \frac{1}{2}(x, Ax) - \nabla(b, x) = Ax - b = 0$ (this uses the fact that $\frac{\partial}{\partial x_i} x = e_i$, the unit vector).

Hence, from both approaches we obtain $Ax = b$,
or in weak form: find a $x_{\mathcal{V}} \in \mathcal{V}$ s.t. $(y, Ax_{\mathcal{V}} - b) = 0 \ \ \forall y \in \mathcal{V}$, which can be written (assuming z solution of $Az = b$) as $(Ay, x_{\mathcal{V}} - z) = 0 \ \ \forall y \in \mathcal{V}$

b. We want to find $x_{\mathcal{V}} \in \mathcal{V}$ such that $(Ay, z - x_{\mathcal{V}}) = 0 \quad \forall y \in \mathcal{V}$ (see part a.1)

Suppose that $\mathcal{V}$ is m-dimensional and V is a matrix whose columns form a basis for $\mathcal{V}$. Then there exists a $\hat{x}_{\mathcal{V}} \in \mathbb{R}^m$ s.t. $x_{\mathcal{V}} = V\hat{x}_{\mathcal{V}}$ and we want to find $\hat{x}_{\mathcal{V}} \in \mathbb{R}^m$ such that

$$(Ay, z - x_{\mathcal{V}}) = (Ay, z - V\hat{x}_{\mathcal{V}}) = (AV\hat{y}, z - V\hat{x}_{\mathcal{V}}) = 0, \quad \forall \hat{y} \in \mathbb{R}^m.$$

Show that that equation can be written as

$$\hat{y}^T V^T AV\hat{x}_{\mathcal{V}} = \hat{y}^T V^T Az, \quad \forall \hat{y} \in \mathbb{R}^m.$$

The latter holds for all $\hat{y} \in \mathbb{R}^m$ if and only if

$$V^T AV\hat{x}_{\mathcal{V}} = V^T Az,$$

and to solve this for $x_{\mathcal{V}}$ we only need to know the product Az, not necessarily z itself.

c. Consider the space $\mathcal{S} = \mathcal{V} \cup \text{span}\{w\}$. We want to find $x_{\mathcal{S}}$ s.t. $(Ay, z - x_{\mathcal{S}}) = 0 \ \forall y \in \mathcal{S}$.

Suppose $[V \ w]$ is a matrix whose columns form a basis for $\mathcal{S}$.

We can write $x_{\mathcal{S}} = [V \ w] \begin{bmatrix} \hat{x}_{\mathcal{V}} \\ \hat{x}_w \end{bmatrix}$ with $\hat{x}_{\mathcal{V}} \in \mathbb{R}^m$ and $\hat{x}_w \in \mathbb{R}$, and analogously $y = [V \ w] \begin{bmatrix} \hat{y}_{\mathcal{V}} \\ \hat{y}_w \end{bmatrix}$.

Show (similar process as in part b) that we can arrive from $x_{\mathcal{S}}$ s.t. $(Ay, z - x_{\mathcal{S}}) = 0 \ \forall y \in \mathcal{S}$ at

$$[V \ w]^T A[V \ w] \begin{bmatrix} \hat{x}_{\mathcal{V}} \\ \hat{x}_w \end{bmatrix} = [V \ w]^T Az \quad \Rightarrow \quad \begin{bmatrix} V^T AV & V^T Aw \\ w^T AV & w^T Aw^T \end{bmatrix} \begin{bmatrix} \hat{x}_{\mathcal{V}} \\ \hat{x}_w \end{bmatrix} = \begin{bmatrix} V^T Az \\ w^T Az \end{bmatrix}$$

where we know that $V^T Aw = 0$ and $w^T AV = 0$ because w is A-orthogonal to $\mathcal{V}$. Hence, we are left with

$$V^T AV\hat{x}_{\mathcal{V}} = V^T Az \quad \text{and} \quad w^T Aw\hat{x}_w = w^T Az,$$

and since we have already solved the first equation when projecting onto $\mathcal{V}$ (in the exercise it was assumed that $x_{\mathcal{V}}$), the scalar equation that still needs to be solved is

$$w^T Aw\hat{x}_w = w^T Az.$$

d. If A not symmetric then $(x, Ay) = (Ax, y)$ does not hold! Therefor, in the computations in part a we need to use $(x, Ay) = (A^T x, y)$. Re-do the directional derivative approach with keeping this in mind.

1. Show $\frac{d}{d\epsilon} J_1(x + \epsilon y)\Big|_{\epsilon=0} = (x - z, A^T y) + (x - z, Ay) = 0$. Hence, minimizing $J_1(x)$ comes down to: find $x_{\mathcal{V}}$ s.t. $(x_{\mathcal{V}} - z, (A + A^T)y) = 0 \ \forall y \in \mathcal{V}$
2. Show $\frac{d}{d\epsilon} J_2(x + \epsilon y)\Big|_{\epsilon=0} = -2(Ay, b - Ax) = 0$. Hence, minimizing $J_2(x)$ comes down to: find $x_{\mathcal{V}}$ s.t. $(Ax_{\mathcal{V}} - b, Ay) = 0 \ \forall y \in \mathcal{V}$
3. Show $\frac{d}{d\epsilon} J_3(x + \epsilon y)\Big|_{\epsilon=0} = \frac{1}{2}(Ay, x_{\mathcal{V}}) + \frac{1}{2}(A^T y, x_{\mathcal{V}}) - (b, y) = 0$. Hence, minimizing $J_3(x)$ comes down to: find $x_{\mathcal{V}}$ s.t. $\frac{1}{2}((A + A^T)y, x_{\mathcal{V}}) - (b, y) = 0 \ \forall y \in \mathcal{V}$

Only for $J_2(x)$ does a non-symmetric matrix A result in the same problem as in case of a symmetric matrix A.
In case of non-symmetric matrix A only for $J_2(x)$ results the minimization process in a solution of $Ax = b$. Minimizing $J_1(x)$ and $J_2(x)$ result in a solution of $\frac{1}{2}(A^T + A)x = b$

Remark:
minimizing J_i (and therewith solving $Ax = b$ with A a symmetric, positive definite matrix in various ways) is analogous to the different ways of solving $\mathcal{A}u = f$ with $\mathcal{A}$ a self-adjoint, coercive operator.

$\mathcal{A}u = f$	$Ax = b$
$\text{argmin}_{v\in\mathcal{V}}(v-u, \mathcal{A}(v-u))$	$\text{argmin}_{x\in\mathcal{V}}(x-z, A(x-z)) = \text{argmin}_{x\in\mathcal{V}} J_1(x)$
$\text{argmin}_{u\in\mathcal{V}}(\mathcal{A}u-f, \mathcal{A}u-f)$	$\text{argmin}_{x\in\mathcal{V}}(b-Ax, b-Ax) = \text{argmin}_{x\in\mathcal{V}} J_2(x)$
$\text{argmin}_{v\in\mathcal{V}}((v, \mathcal{A}v) - 2(f,v))$	$\text{argmin}_{x\in\mathcal{V}}(\frac{1}{2}(x, Ax) - (b,x)) = \text{argmin}_{x\in\mathcal{V}} J_3(x)$