Recall that the derivative of $f$ at $x$ is defined to be the number $l_x$ such that
\[
\lim_{\delta \to 0} \frac{f(x + \delta) - f(x) - l_x \delta}{\delta} = 0
\]
Suppose $f : \mathbb{R}^n \to \mathbb{R}^m$ is differentiable. Then $f'(x)$ is the unique matrix $L_x \in \mathbb{R}^{m \times n}$ which satisfies
\[
\lim_{\delta \to 0, \|\delta\|} \frac{1}{\|\delta\|} (f(x + \delta) - f(x) - L_x \delta) = 0
\]
Hence, while $f : \mathbb{R}^n \to \mathbb{R}^m$, the value of $f'$ at a specific location $x$, $f'(x)$, is a linear operator from $\mathbb{R}^n \to \mathbb{R}^m$. Hence the function $f'$ is a map from $\mathbb{R}^n \to \mathcal{L}(\mathbb{R}^n, \mathbb{R}^m)$.

Similarly, if $f : \mathbb{H}^{n \times n} \to \mathbb{H}^{n \times n}$ is differentiable, then the derivative of $f$ at $x$ is the unique linear operator $L_X : \mathbb{H}^{n \times n} \to \mathbb{H}^{n \times n}$ such that satisfying
\[
\lim_{\Delta \to 0_{n \times n}, \|\Delta\|} \frac{1}{\|\Delta\|} (f(X + \Delta) - f(X) - L_X \Delta) = 0
\]
Derivative of Riccati Lyapunov

\[ f(X) := XDX - A^*X - XA - C \]

Note,

\[
\begin{align*}
f(X + \Delta) &= (X + \Delta) D(X + \Delta) - A^* (X + \Delta) - (X + \Delta) A - C \\
&= XDX - A^*X - XA - C + \Delta DX + XD\Delta - A^* \Delta - \Delta A + \Delta D\Delta \\
&= f(X) - \Delta (A - DX) - (A - DX)^* \Delta + \Delta D\Delta
\end{align*}
\]

For any \( W \in \mathbb{C}^{n \times n} \), let \( \mathcal{L}_W : \mathbb{H}^{n \times n} \to \mathbb{H}^{n \times n} \) be the Lyapunov operator

\[
\mathcal{L}_W(X) := W^*X + XW.
\]

Using this notation,

\[
f(X + \Delta) - f(X) - \mathcal{L}_{A+DX} (\Delta) = \Delta D\Delta
\]

Therefore

\[
\lim_{\Delta \to 0_{n \times n}} \frac{1}{\|\Delta\|} [f(X + \Delta) - f(X) - \mathcal{L}_{A+DX} (\Delta)] = \lim_{\Delta \to 0_{n \times n}} \frac{1}{\|\Delta\|} [\Delta D\Delta] = 0
\]

This means that the Lyaponov operator \( \mathcal{L}_{A+DX} \) is the derivative of the matrix-valued function

\[
XDX - A^*X - XA - C
\]

at \( X \).
Suppose \( f : \mathbb{R} \to \mathbb{R} \) is differentiable. An iterative algorithm to find a root of \( f(x) = 0 \) is

\[
x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)}
\]

Suppose \( f : \mathbb{R}^n \to \mathbb{R}^n \) is differentiable. Then to first order

\[
f(x + \delta) = f(x) + L_x \delta
\]

An iterative algorithm to find a “root” of \( f(x) = 0 \) is obtained by setting the first-order approximation expression to zero, and solving for \( \delta \), yielding the update law (with \( x_{k+1} = x_k + \delta \))

\[
x_{k+1} = x_k - L_{x_k}^{-1}[f(x_k)]
\]

Implicitly, it is written as

\[
L_{x_k} [x_{k+1} - x_k] = -f(x_k) \tag{0.18}
\]

For the Riccati equation

\[
f(X) := XDX - A^*X - XA - C
\]

we have derived that the derivative of \( f \) at \( X \) involves the Lyapunov operator. The iteration in (0.18) appears as

\[
\mathcal{L}_{-A+DX_k} (X_{k+1} - X_k) = -X_kDX_k + A^*X_k + X_kA + C
\]

In detail, this is

\[
-(A - DX_k)^* (X_{k+1} - X_k)(A - DX_k) = -X_kDX_k + A^*X_k + X_kA + C
\]

which simplifies (check) to

\[
(A - DX_k)^* X_{k+1} + X_{k+1} (A - DX_k) = -X_kDX_k - C
\]

Of course, questions about the well-posedness and convergence of the iteration must be addressed in detail.
Comparisons

Theorems 2.1-2.3 in “On Hermitian Solutions of the Symmetric Algebraic Riccati Equation,” by Gohberg, Lancaster and Rodman, SIAM Journal of Control and Optimization, vol. 24, no. 6, Nov. 1986, are the basis for comparing equality and inequality solutions of Riccati equations. The definitions, theorems and main ideas are below.

Suppose that $A \in \mathbb{R}^{n \times n}, C \in \mathbb{R}^{n \times n}, D \in \mathbb{R}^{n \times n}$, with $D = D^T \succeq 0$, $C = C^T$ and $(A, D)$ stabilizable. Consider the matrix equation

$$XDX - A^T X - XA - C = 0 \quad (0.19)$$

**Definition 45** A symmetric solution $X_+ = X_+^T$ of (0.26) is maximal if $X_+ \succeq X$ for any other symmetric solution $X$ of (0.26). Check: Maximal solutions (if they exist) are unique.

**Theorem 46** If there is a symmetric solution to (0.26), then there is a maximal solution $X_+$. The maximal solution satisfies

$$\max_i \text{Re} \lambda_i (A - DX_+) \leq 0$$

**Theorem 47** If there is a symmetric solution to (0.26), then there is a sequence $\{X_j\}_{j=1}^{\infty}$ such that for each $j$, $X_j = X_j^T$ and

- $0 \preceq X_jDX_j - A^T X_j - X_jA - C$
- $X_j \succeq X$ for any $X = X^T$ solving 0.26
- $A - DX_j$ is Hurwitz
- $\lim_{j \to \infty} X_j = X_+$

**Theorem 48** If there are symmetric solutions to (0.26), then for every symmetric $\tilde{C} \succeq C$, there exists symmetric solutions (and hence a maximal solution $\tilde{X}_+$) to

$$XDX - A^T X - XA - \tilde{C} = 0$$

Moreover, the maximal solutions are related by $\tilde{X}_+ \succeq X_+$. 

54
Identities

For any Hermitian matrices \( Y \) and \( \hat{Y} \), it is trivial to verify

\[
Y(A - DY) + (A - DY)^*Y + YDY = Y(A - D\hat{Y}) + (A - D\hat{Y})^*Y + \hat{Y}D\hat{Y} - (Y - \hat{Y})D(Y - \hat{Y}) \tag{0.20}
\]

Let \( X \) denote any Hermitian solution (if one exists) to the ARE,

\[
XDX - A^*X - XA - C = 0 \tag{0.21}
\]

This (the existence of a hermitian solution) is the departure point for all that will eventually follow.

For any hermitian matrix \( Z \)

\[
-C = -XDX + A^*X + XA
= X(A - DX) + (A - DX)^*X + XDX
= X(A - DZ) + (A - DZ)^*X + ZDZ - (X - Z)D(X - Z) \tag{0.22}
\]

This exploits that \( X \) solves the original Riccati equation, and uses the identity (0.20), with \( Y = X, \hat{Y} = Z \).
Take any $\tilde{C} \succeq C$ (which could be $C$, for instance). Also, let $X_0 = X_0^* \succeq 0$ be chosen so that $A - DX_0$ is Hurwitz. Check: Why is this possible?

Iterate (if possible, which we will show is...) as

$$X_{\nu+1} (A - DX_{\nu}) + (A - DX_{\nu})^* X_{\nu+1} = -X_{\nu} DX_{\nu} - \tilde{C} \quad (0.23)$$

Note that by the Hurwitz assumption on $A - DX_0$, at least the first iteration is possible ($\nu = 0$). Also, recall that the iteration is analogous to a first-order algorithm for root finding, $x_{n+1} := x_n - \frac{f(x_n)}{f'(x_n)}$.

Use $Z := X_{\nu}$ in equation (0.22), and subtract from the iteration equation (0.23), leaving

$$\begin{align*}
(X_{\nu+1} - X) (A - DX_{\nu}) + (A - DX_{\nu})^* (X_{\nu+1} - X) \\
= - (X - X_{\nu}) D (X - X_{\nu}) - (\tilde{C} - C) \\
\end{align*} \quad (0.24)$$

which is valid whenever the iteration equation is valid.
Main Induction

Now for the induction: assume \( X_\nu = X_\nu^* \) and \( A - DX_\nu \) is Hurwitz. Note that this is true for \( \nu = 0 \).

Since \( A - DX_\nu \) is Hurwitz, there is a unique solution \( X_{\nu+1} \) to equation (0.23). Moreover, since \( A - DX_\nu \) is Hurwitz, and the right-hand-side of (0.24) is \( \preceq 0 \), it follows that \( X_{\nu+1} - X \succeq 0 \) (note that \( X_0 - X \) is not necessarily positive semidefinite).

Rewrite (0.20) with \( Y = X_{\nu+1} \) and \( \hat{Y} = X_\nu \).

\[
X_{\nu+1} (A - DX_{\nu+1}) + (A - DX_{\nu+1})^* X_{\nu+1} + X_{\nu+1} DX_{\nu+1} = X_{\nu+1} (A - DX_\nu) + (A - DX_\nu)^* X_{\nu+1} + X_\nu DX_\nu - (X_{\nu+1} - X_\nu) D (X_{\nu+1} - X_\nu)
\]

Rewriting,

\[
-C = X_{\nu+1} (A - DX_{\nu+1}) + (A - DX_{\nu+1})^* X_{\nu+1} + X_{\nu+1} DX_{\nu+1} + (X_{\nu+1} - X_\nu) D (X_{\nu+1} - X_\nu)
\]

Writing (0.22) with \( Z = X_{\nu+1} \) gives

\[
-C = X (A - DX_{\nu+1}) + (A - DX_{\nu+1})^* X + X_{\nu+1} DX_{\nu+1} - (X - X_{\nu+1}) D (X - X_{\nu+1})
\]

Subtract these, yielding

\[
-(\tilde{C} - C) = (X_{\nu+1} - X) (A - DX_{\nu+1}) + (A - DX_{\nu+1})^* (X_{\nu+1} - X) + (X_{\nu+1} - X_\nu) D (X_{\nu+1} - X_\nu) + (X - X_{\nu+1}) D (X - X_{\nu+1})
\]  

(0.25)

Now suppose that Re (\( \lambda \)) \( \geq 0 \), and \( v \in \mathbb{C}^n \) such that \( (A - DX_{\nu+1}) v = \lambda v \). Pre and post-multiply (0.25) by \( v^* \) and \( v \) respectively. This gives

\[
-v^* (\tilde{C} - C) v = \underbrace{2 \text{Re}(\lambda)}_{\geq 0} v^* (X_{\nu+1} - X) v + \underbrace{\left\| D^{1/2} (X_{\nu+1} - X_\nu) \right\|_2^2}_{\geq 0} \underbrace{D^{1/2} (X - X_{\nu+1}) v}_{\geq 0}
\]

Hence, both sides are in fact equal to zero, and therefore

\[
v^* (X_{\nu+1} - X_\nu) D (X_{\nu+1} - X_\nu) = 0.
\]
Since $D \geq 0$, it follows that $D(X_{\nu+1} - X_\nu)v = 0$, which means $DX_{\nu+1}v = DX_\nu v$. Hence
\[
\lambda v = (A - DX_{\nu+1})v = (A - DX_\nu)v
\]
But $A - DX_\nu$ is Hurwitz, and $\text{Re}(\lambda) \geq 0$, so $v = 0_n$ as desired. This means $A - DX_{\nu+1}$ is Hurwitz.

Summarizing:

- Suppose there exists a hermitian $X = X^*$ solving
  \[
  XDX - A^*X - XA - C = 0
  \]
- Let $X_0 = X_0^* \succeq 0$ be chosen so $A - DX_0$ is Hurwitz
- Let $\tilde{C}$ be a Hermitian matrix with $\tilde{C} \succeq C$.
- Iteratively define (if possible) a sequence $\{X_\nu\}_\nu^{\infty}$ via
  \[
  X_{\nu+1} (A - DX_\nu) + (A - DX_\nu)^* X_{\nu+1} = -X_\nu DX_\nu - \tilde{C}
  \]

The following is true:

- Given the choice of $X_0$ and $\tilde{C}$, the sequence $\{X_\nu\}_\nu^{\infty}$ is well-defined, unique, and has $X_\nu = X_\nu^*$.
- For each $\nu \geq 0$, $A - DX_\nu$ is Hurwitz
- For each $\nu \geq 1$, $X_\nu \succeq X$.

It remains to show that for $\nu \geq 1$, $X_{\nu+1} \preceq X_\nu$ (although it is not necessarily true that $X_1 \preceq X_0$).
For $\nu \geq 1$, apply (0.20) with $Y = X_\nu$ and $\hat{Y} = X_{\nu-1}$, and substitute the iteration (shifted back one step) equation (0.23) giving

$$X_\nu (A - DX_\nu) + (A - DX_\nu)^* X_\nu + X_\nu DX_\nu = -\tilde{C} - (X_\nu - X_{\nu-1}) D (X_\nu - X_{\nu-1})$$

The iteration equation is

$$X_{\nu+1} (A - DX_\nu) + (A - DX_\nu)^* X_{\nu+1} = -X_\nu DX_\nu - \tilde{C}$$

Subtracting leaves

$$(X_\nu - X_{\nu+1}) (A - DX_\nu) + (A - DX_\nu)^* (X_\nu - X_{\nu+1}) = (X_\nu - X_{\nu-1}) D (X_\nu - X_{\nu-1})$$

The fact that $A - DX_\nu$ is Hurwitz and $D \succeq 0$ implies that $X_\nu - X_{\nu+1} \succeq 0$.

Hence, we have shown that for all $\nu \geq 1$,

$$X \preceq X_{\nu+1} \preceq X_\nu$$
Monotonic Matrix Sequences

Every nonincreasing sequence of real numbers which is bounded below has a limit. Specifically, if \( \{x_k\}_{k=0}^{\infty} \) is a sequence of real numbers, and there is a real number \( \gamma \) such that \( x_k \geq \gamma \) and \( x_{k+1} \leq x_k \) for all \( k \), then there is a real number \( x \) such that

\[
\lim_{k \to \infty} x_k = x
\]

A similar result holds for symmetric matrices: Specifically, if \( \{X_k\}_{k=0}^{\infty} \) is a sequence of real, symmetric matrices, and there is a real symmetric matrix \( \Gamma \) such that \( X_k \succeq \Gamma \) and \( X_{k+1} \preceq X_k \) for all \( k \), then there is a real symmetric matrix \( X \) such that

\[
\lim_{k \to \infty} X_k = X
\]

This is proven by first considering the sequence \( \{e_i^T X_k e_i\}_{k=0}^{\infty} \), which shows that all of the diagonal entries of the sequence have limits. Then look at \( \{(e_i + e_j)^T X_k (e_i + e_j)\}_{k=0}^{\infty} \) to conclude convergence of off-diagonal terms.
The sequence \( \{X_\nu\}_{\nu=0}^\infty \), generated by (0.23), has a limit
\[
\lim_{\nu \to \infty} X_\nu =: \tilde{X}_+
\]
Remember that \( \tilde{C} \) was any hermitian matrix \( \succeq C \).

Facts:

- Since each \( A - DX_\nu \) is Hurwitz, in passing to the limit we may get eigenvalues with zero real-parts, hence
\[
\max_i \text{Re} \lambda_i (A - DX_+) \leq 0
\]

- Since each \( X_\nu = X_\nu^* \), it follows that \( \tilde{X}^*_+ = \tilde{X}_+ \).

- Since each \( X_\nu \succeq X \) (where \( X \) is any hermitian solution to the equality \( XDX - A^*X - XA - C = 0 \), equation (0.21)), it follows by passing to the limit that \( \tilde{X}_+ \succeq X \).

- The iteration equation is
\[
X_{\nu+1} (A - DX_\nu) + (A - DX_\nu)^* X_{\nu+1} = -X_\nu DX_\nu - \tilde{C}
\]
Taking limits yields
\[
\tilde{X}_+ (A - D\tilde{X}_+) + (A - D\tilde{X}_+)^* \tilde{X}_+ = -\tilde{X}_+ D\tilde{X}_+ - \tilde{C}
\]
which has a few cancellations, leaving
\[
\tilde{X}_+ A + A^* \tilde{X}_+ - \tilde{X}_+ D\tilde{X}_+ + \tilde{C} = 0
\]

- Finally, if \( X_+ \) denotes the above limit when \( \tilde{C} = C \), it follows that the above ideas hold, so \( X_+ \succeq X \), for all solutions \( X \).

- And really finally, since \( X_+ \) is a hermitian solution to (0.21) with \( C \), the properties of \( \tilde{X}_+ \) imply that \( \tilde{X}_+ \succeq X_+ \).
Analogies with Scalar Case

\[ dx^2 - 2ax - c = 0 \]

<table>
<thead>
<tr>
<th>Matrix Statement</th>
<th>Scalar Specialization</th>
</tr>
</thead>
<tbody>
<tr>
<td>( D = D^* \succeq 0 )</td>
<td>( d \in \mathbb{R}, \text{ with } d \geq 0 )</td>
</tr>
<tr>
<td>( (A, D) ) stabilizable</td>
<td>( a &lt; 0 ) or ( d &gt; 0 )</td>
</tr>
<tr>
<td>existence of Hermitian solution</td>
<td>( a^2 + cd \geq 0 )</td>
</tr>
<tr>
<td>Maximal solution ( X_+ )</td>
<td>( x_+ = \frac{a + \sqrt{a^2 + cd}}{d} ) for ( d &gt; 0 )</td>
</tr>
<tr>
<td>Maximal solution ( X_+ )</td>
<td>( x_+ = \frac{c}{-2a} ) for ( d = 0 )</td>
</tr>
<tr>
<td>( A - DX_+ )</td>
<td>(-\sqrt{a^2 + cd} ) for ( d &gt; 0 )</td>
</tr>
<tr>
<td>( A - DX_0 ) Hurwitz</td>
<td>( a ) for ( d = 0 )</td>
</tr>
<tr>
<td>( A - DX_0 ) Hurwitz</td>
<td>( x_0 &gt; \frac{a}{d} ) for ( d &gt; 0 )</td>
</tr>
<tr>
<td>( A - DX ) Hurwitz</td>
<td>( a &lt; 0 ) for ( d = 0 )</td>
</tr>
<tr>
<td>( A - DX ) Hurwitz</td>
<td>( x &gt; \frac{a}{d} ) for ( d &gt; 0 )</td>
</tr>
<tr>
<td>( A - DX ) Hurwitz</td>
<td>( a &lt; 0 ) for ( d = 0 )</td>
</tr>
</tbody>
</table>

If \( d = 0 \), then the graph of \(-2ax - c\) (with \( a < 0 \)) is a straight line with a positive slope. There is one (and only one) real solution at \( x = -\frac{c}{2a} \).

If \( d > 0 \), then the graph of \( dx^2 - 2ax - c \) is a upward directed parabola. The minimum occurs at \( \frac{a}{d} \). There are either 0, 1, or 2 real solutions to \( dx^2 - 2ax - c = 0 \).

- If \( a^2 + cd < 0 \), then there are no solutions (the minimum is positive)
- If \( a^2 + cd = 0 \), then there is one solution, at \( x = \frac{a}{d} \), which is where the minimum occurs.
- If \( a^2 + cd > 0 \), then there are two solutions, at \( x_- = \frac{a}{d} - \frac{\sqrt{a^2 + cd}}{d} \) and \( x_+ = \frac{a}{d} + \frac{\sqrt{a^2 + cd}}{d} \) which are equidistant from the minimum \( \frac{a}{d} \).