Kalman Filtering with Inequality Constraints for Turbofan Engine Health Estimation

Dan Simon *  
d.j.simon@csuohio.edu  
Cleveland State University  
Stilwell Hall Room 332  
1960 East 24th Street  
Cleveland, OH 44115

Donald L. Simon
US Army Research Laboratory
NASA Glenn Research Center
Mail Stop 77-1
21000 Brookpark Road
Cleveland, OH 44135

Abstract

Kalman filters are often used to estimate the state variables of a dynamic system. However, in the application of Kalman filters some known signal information is often either ignored or dealt with heuristically. For instance, state variable constraints (which may be based on physical considerations) are often neglected because they do not fit easily into the structure of the Kalman filter. This paper develops two analytic methods of incorporating state variable inequality constraints in the Kalman filter. The first method is a general technique of using hard constraints to enforce inequalities on the state variable estimates. The resultant filter is a combination of a standard Kalman filter and a quadratic programming problem. The second method uses soft constraints to estimate state variables that are known to vary slowly with time. (Soft constraints are constraints that are required to be approximately satisfied rather than exactly satisfied.) The incorporation of state variable constraints increases the computational effort of the filter but significantly improves its estimation accuracy. The improvement is proven theoretically and shown via simulation.

*Corresponding author. This work was supported in part by a NASA/ASEE Summer Faculty Fellowship.
results. The use of the algorithm is demonstrated on a linearized simulation of a turbofan engine to estimate health parameters. The turbofan engine model contains 16 state variables, 12 measurements, and 8 component health parameters. It is shown that the new algorithms provide improved performance in this example over unconstrained Kalman filtering.

Key Words – Kalman Filter, State Constraints, Estimation, Quadratic Programming, Gas Turbine Engines.

1 Introduction

For linear dynamic systems with white process and measurement noise, the Kalman filter is known to be an optimal estimator. However, in the application of Kalman filters there is often known model or signal information that is either ignored or dealt with heuristically [1]. This paper presents two ways to generalize the Kalman filter in such a way that known inequality constraints among the state variables are satisfied by the state variable estimates.

The first method presented here for enforcing inequality constraints on the state variable estimates uses hard constraints. It is based on a generalization of the approach presented in [2], which dealt with the incorporation of state variable equality constraints in the Kalman filter. Inequality constraints are inherently more complicated than equality constraints, but standard quadratic programming results can be used to solve the Kalman filter problem with inequality constraints. At each time step of the constrained Kalman filter, we solve a quadratic programming problem to obtain the constrained state estimate. A family of constrained state estimates is obtained, where the weighting matrix of the quadratic programming problem determines which family member forms the desired solution. It is stated in this paper, on the basis of [2], that the constrained estimate has several important properties. The constrained state estimate is unbiased and has a smaller error covariance than the unconstrained estimate. We show which member of all possible constrained solutions has the smallest error covariance. We also show the one particular member that is always (i.e., at each time step) closer to the true state than the unconstrained
estimate.

The second method for enforcing inequality constraints uses soft constraints via a penalty term in the Kalman filter optimization problem. This prevents the state estimate from changing too rapidly. It essentially smooths the unconstrained Kalman filter estimate when the state variables are known to vary slowly with time. It is shown that the constrained state estimate is unbiased, approaches the unconstrained estimate as time approaches infinity, and (under certain special conditions) is equal to the running average of the unconstrained estimate.

The application considered in this paper is turbofan engine health parameter estimation [3]. The performance of gas turbine engines deteriorates over time. This deterioration reduces the fuel economy of the engine. Airlines periodically collect engine data in order to evaluate the health of the engine and its components. The health evaluation is then used to determine maintenance schedules. Reliable health evaluations are used to anticipate future maintenance needs. This offers the benefits of improved safety and reduced operating costs. The money-saving potential of such health evaluations is substantial, but only if the evaluations are reliable. The data used to perform health evaluations are typically collected during flight and later transferred to ground-based computers for post-flight analysis. Data are collected each flight at the same engine operating points and corrected to account for variability in ambient conditions. Typically, data are collected for a period of about 3 seconds at a rate of about 10 or 20 Hz. Various algorithms have been proposed to estimate engine health parameters, such as weighted least squares [4], expert systems [5], Kalman filters [6], neural networks [6], and genetic algorithms [7].

This paper applies constrained Kalman filtering to estimate engine component efficiencies and flow capacities, which are referred to as health parameters. We can use our knowledge of the physics of the turbofan engine in order to obtain a dynamic model [8, 9]. The health parameters that we try to estimate can be modelled as slowly varying biases. The state vector of the dynamic model is augmented to include the health parameters, which are then estimated with a Kalman filter [10]. The model formulation in this paper is similar to previous NASA work [11]. However, [11] was limited to a 3-state dynamic model and 2 health parameters, whereas this
present work includes a more complete 16-state model and 8 health parameters. In addition, we have some a priori knowledge of the engine’s health parameters: we know that they never improve. Engine health always degrades over time, and we can incorporate this information into state constraints to improve our health parameter estimation. (This is assuming that no maintenance or engine overhaul is performed.) This is similar to the probabilistic approach to turbofan prognostics proposed in [12].

It should be emphasized that in this paper we are confining the problem to the estimation of engine health parameters in the presence of degradation only. There are specific engine cases that can result in abrupt shifts in filter estimates, possibly even indicating an apparent improvement in some engine components. An actual engine performance monitoring system would need to include additional logic to detect and isolate such faults.

Section 2 presents a discussion of the standard discrete time Kalman filter. Some important properties of the Kalman filter that will be used later in this paper are also reviewed. Section 3 generalizes the results of [2] to hard inequality constraints. This inequality-constrained Kalman filter has several attractive theoretical properties, including state variable estimates that are unbiased, an estimation error variance smaller than the unconstrained filter, and a time-domain estimation error that is always smaller than the unconstrained estimation error. Section 4 extends the standard Kalman filter in a different way for those cases where it is known that the state variables change slowly with time. This constraint is enforced by finding a new state estimate that is “close” to the unconstrained estimate in some sense, but that is slowly time varying. It is shown that this new estimate is unbiased, approaches the unconstrained estimate as time goes to infinity, and (under certain conditions) is equal to the running average of the unconstrained estimate.

Section 5 discusses the problem of turbofan health parameter estimation, along with the dynamic model that we used in our simulation experiments. Although the health parameters are not state variables of the model, it is shown how the dynamic model can be augmented in such a way that a Kalman filter can estimate the health parameters [10, 11]. We then show how this problem can be expressed in such a way to be compatible with the constraints discussed in the preceding sections.
Section 6 presents some simulation results based on a turbofan model linearized around a known operating point. We show that the Kalman filter can estimate health parameters with an average error of less than 0.2%, and the constrained Kalman filters perform better than the unconstrained filter. Section 7 presents some concluding remarks and suggestions for further work.

2 Kalman Filtering

This section reviews standard (unconstrained) state estimation via the Kalman filter and some important properties of the filter that will be used later in this paper. The results and notation are taken from [13]. Consider the discrete linear time-invariant system given by

\[
x_{k+1} = Ax_k + Bu_k + w_k
\]

\[
y_k = Cx_k + e_k
\]

where \( k \) is the time index, \( x \) is the state vector, \( u \) is the known control input, \( y \) is the measurement, and \( \{w_k\} \) and \( \{e_k\} \) are noise input sequences. The problem is to find an estimate \( \hat{x}_{k+1} \) of \( x_{k+1} \) given the measurements \( \{y_0, y_1, \cdots, y_k\} \). We will use the symbol \( Y_k \) to denote the column vector that contains the measurements \( \{y_0, y_1, \cdots, y_k\} \). We assume that the following standard conditions are satisfied.

\[
E[x_0] = \bar{x}_0
\]

\[
E[w_k] = E[e_k] = 0
\]

\[
E[(x_0 - \bar{x}_0)(x_0 - \bar{x}_0)^T] = \Sigma_0
\]

\[
E[w_kw_m^T] = Q\delta_{km}
\]

\[
E[e_ke_m^T] = R\delta_{km}
\]

\[
E[w_ke_m^T] = E[x_ke_m^T] = E[x_kw_m^T] = 0
\]

where \( E[\cdot] \) is the expectation operator, \( \bar{x} \) is the expected value of \( x \), and \( \delta_{km} \) is the Kronecker delta function (\( \delta_{km} = 1 \) if \( k = m, 0 \) otherwise). \( Q \) and \( R \) are positive semidefinite covariance matrices. The Kalman filter equations are given by

\[
K_k = A\Sigma_kC^T(C\Sigma_kC^T + R)^{-1}
\]
\[ \hat{x}_{k+1} = A\hat{x}_k + Bu_k + K_k(y_k - C\hat{x}_k) \]  
\[ \Sigma_{k+1} = (A\Sigma_k - K_kCG)A^T + Q \]  
(9)  
(10)

where the filter is initialized with \( \hat{x}_0 = \bar{x}_0 \), and \( \Sigma_0 \) given above. It can be shown [13] that the Kalman filter has several attractive properties. For instance, if \( x_0, \{w_k\}, \) and \( \{e_k\} \) are jointly gaussian, the Kalman filter estimate \( \hat{x}_{k+1} \) is the conditional mean of \( x_{k+1} \) given the measurements \( Y_k \); i.e., \( \hat{x}_{k+1} = E[x_{k+1}|Y_k] \). Even if \( x_0, \{w_k\}, \) and \( \{e_k\} \) are not jointly gaussian, the Kalman filter estimate is the best affine estimator given the measurements \( Y_k \); i.e., of all estimates of \( x_{k+1} \) that are of the form \( FY_k + g \) (where \( F \) is a fixed matrix and \( g \) is a fixed vector), the Kalman filter estimate is the one that minimizes the variance of the estimation error. It can be shown [13, pp. 92 ff.] that the Kalman filter estimate (i.e., the minimum variance estimate) can be given by

\[ \hat{x}_{k+1} = \bar{x}_{k+1} \equiv \bar{x}_{k+1} + \Sigma_{xy}\Sigma_{yy}^{-1}(Y_k - \bar{Y}_k) \]  
(11)

where \( \bar{x}_{k+1} \) is the mean of \( x_{k+1} \), \( \Sigma_{xy} \) is the variance matrix of \( x_{k+1} \) and \( Y_k \), \( \Sigma_{yy} \) is the covariance matrix of \( Y_k \), and \( \bar{x}_{k+1} \) is the conditional mean of \( x_{k+1} \) given the measurements \( Y_k \). In addition, from [13, p. 93] we know that the Kalman filter estimate \( \hat{x}_{k+1} \) and \( Y_k \) are jointly gaussian, in which case \( \hat{x}_{k+1} \) is conditionally gaussian given \( Y_k \). The conditional probability density function of \( x_{k+1} \) given \( Y_k \) is

\[ P(x|Y) = \frac{\exp[-(x - \bar{x})^T\Sigma^{-1}(x - \bar{x})/2]}{(2\pi)^{n/2}|\Sigma|^{1/2}} \]  
(12)

where \( n \) is the dimension of \( x \) and

\[ \Sigma = \Sigma_{xx} - \Sigma_{xy}\Sigma_{yy}^{-1}\Sigma_{yx} \]  
(13)

The Kalman filter estimate is that value of \( x \) that maximizes the conditional probability density function \( P(x|Y) \), and \( \Sigma \) is the covariance of the Kalman filter estimation error.
3 Kalman Filtering with Hard Inequality Constraints

This section extends the well known results of the previous section to cases where there are known linear inequality constraints among the state components. Also, several important properties of the constrained filter are discussed. Consider the dynamic system of (1) where we are given the additional constraint

$$Dx_k \leq d_k$$  \hspace{1cm} (14)

where $D$ is a known $s \times n$ constant matrix, $s$ is the number of constraints, $n$ is the number of state variables, and $s \leq n$. It is assumed in this paper that $D$ is full rank, i.e., that $D$ has rank $s$. This is an easily satisfied assumption. If $D$ is not full rank that means we have redundant state constraints. In that case we can simply remove linearly dependent rows from $D$ (i.e., remove redundant state constraints) until $D$ is full rank. Three different approaches to the constrained state estimation problem are given in this section. The time index $k$ is omitted in the remainder of this section for ease of notation.

3.1 The Maximum Probability Method

In this section we derive the constrained Kalman filtering problem by using a maximum probability method. From [13, pp. 93 ff.] we know that the Kalman filter estimate is that value of $x$ that maximizes the conditional probability density function $P(x|Y)$, which is given in (12). The constrained Kalman filter can be derived by finding an estimate $\tilde{x}$ such that the conditional probability $P(\tilde{x}|Y)$ is maximized and $\tilde{x}$ satisfies the constraint (14). Maximizing $P(\tilde{x}|Y)$ is the same as maximizing its natural logarithm. So the problem we want to solve can be given by

$$\max \ln P(\tilde{x}|Y) \quad \Rightarrow \quad \min_{\tilde{x}} (\tilde{x} - \bar{x})^T \Sigma^{-1} (\tilde{x} - \bar{x})$$  \hspace{1cm} (15)

such that $D\tilde{x} \leq d$
Using the fact that the unconstrained state estimate $\hat{x} = \bar{x}$ (the conditional mean of $x$), we rewrite the above equation as

$$\min_{\bar{x}} (\bar{x}^T \Sigma^{-1} \bar{x} - 2\bar{x}^T \Sigma^{-1} \bar{x}) \text{ such that } D\bar{x} \leq d$$

(16)

Note that this problem statement depends on the conditional gaussian nature of $\hat{x}$, which in turn depends on the gaussian nature of $x_0$, $\{w_k\}$, and $\{e_k\}$ in (1).

### 3.2 The Mean Square Method

In this section we derive the constrained Kalman filtering problem by using a mean square minimization method. We seek to minimize the conditional mean square error subject to the state constraints.

$$\min_{\bar{x}} E(\|x - \bar{x}\|^2|Y) \text{ such that } D\bar{x} \leq d$$

(17)

where $\| \cdot \|$ denotes the vector two-norm. If we assume that $x$ and $Y$ are jointly gaussian, the mean square error can be written as

$$E(\|x - \hat{x}\|^2|Y) = \int (x - \hat{x})^T (x - \hat{x}) P(x|Y) dx$$

(18)

$$= \int x^T x P(x|Y) dx - 2\hat{x}^T \int x P(x|Y) dx + \hat{x}^T \hat{x}$$

(19)

Noting that the Kalman filter estimate is the conditional mean of $x$, i.e.,

$$\hat{x} = \int x P(x|Y) dx$$

(20)

we formulate the first order conditions necessary for a minimum as

$$\min_{\bar{x}} (\bar{x}^T \bar{x} - 2\bar{x}^T \hat{x}) \text{ such that } D\bar{x} \leq d$$

(21)

Again, this problem statement depends on the conditional gaussian nature of $\hat{x}$, which in turn depends on the gaussian nature of $x_0$, $\{w_k\}$, and $\{e_k\}$ in (1).

### 3.3 The Projection Method

In this section we derive the constrained Kalman filtering problem by directly projecting the unconstrained state estimate $\hat{x}$ onto the constraint surface. That is, we
solve the problem
\[
\min_{\hat{x}} (\hat{x} - \hat{x})^TW(\hat{x} - \hat{x}) \text{ such that } D\hat{x} \leq d
\]  \hspace{1cm} (22)

where \( W \) is any symmetric positive definite weighting matrix. This problem can be rewritten as
\[
\min_{\hat{x}} (\hat{x}^TW\hat{x} - 2\hat{x}^T\hat{x}) \text{ such that } D\hat{x} \leq d
\]  \hspace{1cm} (23)

The constrained estimation problems derived by the maximum probability method (16) and the mean square method (21) can be obtained from this equation by setting \( W = \Sigma^{-1} \) and \( W = I \) respectively. Note that this derivation of the constrained estimation problem does not depend on the conditional gaussian nature of \( \hat{x} \); i.e., \( x_0, \{w_k\}, \text{ and } \{e_k\} \) in (1) are not assumed to be gaussian.

### 3.4 The Solution of the Constrained State Estimation Problem

The problem defined by (23) is known as a quadratic programming problem [14, 15]. There are many algorithms for solving quadratic programming problems, almost all of which fall in the category known as active set methods. An active set method uses the fact that it is only those constraints that are active at the solution of the problem that are significant in the optimality conditions. Assume that \( t \) of the \( s \) inequality constraints are active at the solution of (23), and denote by \( \hat{D} \) and \( \hat{d} \) the \( t \) rows of \( D \) and \( t \) elements of \( d \) corresponding to the active constraints. If the correct set of active constraints was known \textit{a priori} then the solution of (23) would also be a solution of the equality-constrained problem
\[
\min_{\hat{x}} (\hat{x}^TW\hat{x} - 2\hat{x}^T\hat{x}) \text{ such that } \hat{D}\hat{x} = \hat{d}
\]  \hspace{1cm} (24)

This shows that the inequality constrained problem defined by (23) is equivalent to the equality-constrained problem defined by (24). The equality-constrained problem was discussed in [2], and so those results can be used to investigate the properties of the inequality-constrained problem.
3.5 Properties of the Constrained State Estimate

In this section we examine some of the statistical properties of the constrained Kalman filter. We use \( \hat{x} \) to denote the state estimate of the unconstrained Kalman filter, and \( \tilde{x} \) to denote the state estimate of the constrained Kalman filter as given by (23), recalling that (16) and (21) are special cases of (23).

**Theorem 1** The solution \( \tilde{x} \) of the constrained state estimation problem given by (23) is an unbiased state estimator for the system (1) for any symmetric positive definite weighting matrix \( W \). That is,

\[
E(\tilde{x}) = E(x)
\]  

(25)

**Theorem 2** The solution \( \tilde{x} \) of the constrained state estimation problem given by (23) with \( W = \Sigma^{-1} \), where \( \Sigma \) is the covariance of the unconstrained estimate given in (10) and (13), has an error covariance that is less than or equal to that of the unconstrained state estimate. That is,

\[
\text{Cov}(x - \tilde{x}) \leq \text{Cov}(x - \hat{x})
\]  

(26)

At first this seems counterintuitive, since the standard Kalman filter is by definition the minimum variance filter. However, we have changed the problem by introducing state variable constraints. Therefore, the standard Kalman filter is no longer the minimum variance filter, and we can do better with the constrained Kalman filter.

**Theorem 3** Among all the constrained Kalman filters resulting from the solution of (23), the filter that uses \( W = \Sigma^{-1} \) has the smallest estimation error covariance. That is,

\[
\text{Cov}(\tilde{x}_{\Sigma^{-1}}) \leq \text{Cov}(\tilde{x}_W) \text{ for all } W
\]  

(27)

**Theorem 4** The solution \( \tilde{x} \) of the constrained state estimation problem given by (23) with \( W = I \) satisfies the inequality

\[
\|x_k - \tilde{x}_k\| \leq \|x_k - \hat{x}_k\| \text{ for all } k
\]  

(28)

where \( \| \cdot \| \) is the vector two-norm and \( \hat{x} \) is the unconstrained Kalman filter estimate.
Theorem 5  The error of the solution $\hat{x}$ of the constrained state estimation problem given by (23) with $W = I$ is smaller than the unconstrained estimation error in the sense that

$$\text{Tr}[	ext{Cov}(\hat{x})] \leq \text{Tr}[	ext{Cov}(\tilde{x})]$$

(29)

where $\text{Tr}[\cdot]$ indicates the trace of a matrix, and $\text{Cov}(\cdot)$ indicates the covariance matrix of a random vector.

The above theorems all follow from the equivalence of (23) and (24), and the proofs presented in [2]. We note that if any of the $s$ constraints are active at the solution of (23), then strict inequalities hold in the statements of Theorems 2–5. The only time that equalities hold in the theorems is if there are no active constraints at the solution of (23); that is, if the unconstrained Kalman filter satisfies the inequality constraints.

4 Kalman Filtering with Soft Inequality Constraints

In this section we are interested in obtaining a Kalman filter-based state estimate for state variables which we know a priori vary slowly with time. Since we are concerned with using the Kalman filter as a parameter estimator, we will assume for this problem that the $A$ matrix in (1) is the identity matrix and the $B$ matrix is zero.

With this in mind, we can use the results of the previous section, especially (22), to formulate a Kalman filter-based estimate as follows

$$\min_{\tilde{x}_k} (\tilde{x}_k - \hat{x}_k)^T W (\tilde{x}_k - \hat{x}_k) \text{ such that } \tilde{x}(i) \text{ varies slowly}$$

(30)

where, as before, $W$ is a constant symmetric positive definite weighting matrix. This is a type of regularization; that is, some additional structure is incorporated into the Kalman filter estimate [16, 17, 18]. The above problem can be formulated as

$$\min_{\tilde{x}_k} [(\tilde{x}_k - \hat{x}_k)^T W (\tilde{x}_k - \hat{x}_k) + (\tilde{x}_k - \tilde{x}_{k-1})^T V_k (\tilde{x}_k - \tilde{x}_{k-1})]$$

(31)
where $V_k$ is a (possibly time-varying) symmetric positive definite weighting matrix that balances the desire for a close approximation to $\hat{x}$ and smooth estimate $\tilde{x}$. The solution to the above problem is

\[
\tilde{x}_0 = E[x_0] \\
\tilde{x}_k = (W + V_k)^{-1}(W\hat{x}_k + V_k\tilde{x}_{k-1})
\]

Since $W$ and $V_k$ are both positive definite, we know that $(W + V_k)^{-1}$ exists.

**Theorem 6** Assume (as stated above) that $A = I$ and $B = 0$ in (1). Then the solution $\tilde{x}$ of the constrained state estimation problem given by (32) is an unbiased state estimator for the system (1) for any symmetric positive definite weighting matrices $W$ and $V_k$. That is,

\[E(\tilde{x}) = E(x)\]  

**Proof:** The theorem can be proven by induction. Since $A = I$ and $B = 0$ we know that $E[x_k] = \bar{x}_0$ for all $k$. We therefore know from (32) that $\tilde{x}_0 = \bar{x}_0$. From (32) with $k = 0$ we see that $E[\tilde{x}_1] = \bar{x}_0$. We repeat this process to show that $E[\tilde{x}_k] = E[\tilde{x}_k] = \bar{x}_0$ for all $k$.

QED

**Theorem 7** Assume (as stated above) that $A = I$ and $B = 0$ in (1). Further assume that $w_k = 0$ in (1) (since we are trying to estimate constant parameters). Then the constrained state estimate $\tilde{x}$ approaches the unconstrained estimate $\hat{x}$ in the limit as time goes to infinity. That is,

\[
\lim_{k \to \infty} \tilde{x}_k = \lim_{k \to \infty} \hat{x}_k
\]

**Proof:** We see from (8)–(10) that, under the conditions stated here, $K_k \to 0$ as $k \to \infty$. Therefore $\hat{x}_k$ approaches a constant value as $k \to \infty$. From (32) we see that, in steady state

\[
\hat{x} = (W + V_k)^{-1}(W\hat{x} + V_k\tilde{x})
\]

\[\implies \hat{x} = [I - (W + V_k)^{-1}V_k]^{-1}(W + V_k)^{-1}W\hat{x}
\]

\[= (I + W^{-1}V_k)(W + V_k)^{-1}W\hat{x}
\]
where the last equality follows from the matrix inversion lemma. Premultiplying both sides of the above equation by $W$ we obtain $W \hat{x} = W \hat{\hat{x}}$, so if $W$ is invertible (which it is, since we are assuming in this section that $W$ is positive definite), we obtain $\hat{x} = \hat{\hat{x}}$ (in steady state). Note that the theorem is true even if $V_k$ does not approach a steady state value as $k \to \infty$.

QED

**Theorem 8** If $V_k = (k-1)W$ in (32) then $\hat{x}_k$ is the running average of $\hat{\hat{x}}_k$.

**Proof:** The running average of $\hat{x}_k$ is defined as

$$X_k = \frac{1}{k} \sum_{i=1}^{k} \hat{x}_i$$

which implies that

$$X_{k+1} = \frac{1}{k+1} (\hat{x}_{k+1} + kX_k)$$

Now if $V_k = (k-1)W$ then (32) shows that

$$\hat{x}_{k+1} = [(k+1)W]^{-1} (W\hat{x}_{k+1} + kW\hat{x}_k)$$

$$= \frac{1}{k+1} (\hat{x}_{k+1} + k\hat{x}_k)$$

which is exactly the running average shown in (37).

QED

## 5 Turbofan Engine Health Monitoring

Figure 1 shows a schematic representation of a turbofan engine. A single inlet supplies airflow to the fan. Air leaving the fan separates into two streams: one stream passes through the engine core, and the other stream passes through the annular bypass duct. The fan is driven by the low pressure turbine. The air passing through the engine core moves through the compressor, which is driven by the high pressure turbine. Fuel is injected in the main combustor and burned to produce hot gas for driving the turbines. The two air streams combine in the augmentor
duct, where additional fuel is added to further increase the air temperature. The air leaves the augmentor through the nozzle, which has a variable cross section area.

Various turbofan simulation packages have been proposed over the years [19, 20, 21]. This model is based on a gas turbine engine simulation software package called DIGTEM (Digital Turbofan Engine Model) [8, 22]. DIGTEM is written in Fortran and includes 16 state variables. It uses a backward difference integration scheme because the turbofan model contains time constants that differ by up to four orders of magnitude.

The nonlinear equations used in DIGTEM can be found in [8, 9]. The time-invariant equations can be summarized as follows.

\[
\begin{align*}
\dot{x} &= f(x, u, p) + w_1(t) \\
y &= g(x, u, p) + e(t)
\end{align*}
\]  

(39)

\(x\) is the 16-element state vector, \(u\) is the 6-element control vector, \(p\) is the 8-element vector of health parameters, and \(y\) is the 12-element vector of measurements. The noise term \(w_1(t)\) represents inaccuracies in the model, and \(e(t)\) represents measurement noise. The elements in these vectors are summarized in Tables 1–4, along with their values at the nominal operating point \((x_0, u_0, p_0, y_0)\) considered in this paper. Table 4 also shows typical signal-to-noise ratios for the measurements, based on NASA experience and previously published data [23]. Sensor dynamics are assumed to be high enough bandwidth that they can be ignored in the dynamic equations [23].

Equation (39) can be linearized about the nominal operating point by using the first order approximation of the Taylor series expansion

\[
\begin{align*}
f(x, u, p) &\approx f(x_0, u_0, p_0) + \frac{\partial f(\cdot)}{\partial x}(x - x_0) + \frac{\partial f(\cdot)}{\partial u}(u - u_0) + \frac{\partial f(\cdot)}{\partial p}(p - p_0) + w_1(t) \\
g(x, u, p) &\approx g(x_0, u_0, p_0) + \frac{\partial g(\cdot)}{\partial x}(x - x_0) + \frac{\partial g(\cdot)}{\partial u}(u - u_0) + \frac{\partial g(\cdot)}{\partial p}(p - p_0) + e(t)
\end{align*}
\]  

(40)

Therefore, a linear small signal system model can be defined for small excursions from the nominal operating point.
\[
\delta \dot{x} \equiv \dot{x} - \dot{x}_0 = A_1 \delta x + B \delta u + A_2 \delta p + w_1(t) \tag{41}
\]
\[
\delta y \equiv y - y_0 = C_1 \delta x + D \delta u + C_2 \delta p + e(t)
\]

We note that
\[
A_1 = \frac{\partial f}{\partial x}
\]
\[
A_{1(i,j)} \approx \frac{\Delta \dot{x}(i)}{\Delta x(j)}
\]

Similar equations hold for the \(A_2\), \(C_1\), and \(C_2\) matrices. We obtained numerical approximations to the \(A_1\), \(A_2\), \(C_1\), and \(C_2\) matrices by varying \(x\) and \(p\) from their nominal values (one element at a time) and recording the new \(\dot{x}\) and \(y\) vectors in DIGTEM.

Turbofan engine health monitoring is typically a two-step process \([3]\). In the first step, engine data is collected each flight at the same engine operating points and corrected to account for variability in ambient conditions. Data are typically collected for a period of about 3 seconds per flight at a rate of about 10 or 20 Hz. In the second step, the data are transferred to ground-based computers for post-flight analysis to determine engine health.

The goal of our turbofan engine health monitoring problem is to obtain an accurate estimate of \(\delta p\), which varies slowly with time. We therefore assume that \(\delta p\) is constant between measurement times. We also assume that the control input is perfectly known, so \(\delta u = 0\). This gives us the following equivalent discrete time system \([24\text{, pp. 90 ff.}]\).

\[
\delta x_{k+1} = A_{1d} \delta x_k + A_{2d} \delta p_k + w_{1k} \tag{43}
\]
\[
\delta y_k = C_1 \delta x_k + C_2 \delta p_k + e_k
\]

where \(A_{1d} = \exp(A_1 T)\) and \(A_{2d} = A_1^{-1}(A_{1d} - I)A_2\) (assuming that \(A_1\) is invertible, which it is in our problem). We next augment the state vector with the health parameter vector \([11]\) to obtain the system equation.
where $w_{2k}$ is a small noise term (uncorrelated with $w_{1k}$) that represents model uncertainty and allows the Kalman filter to estimate time-varying health parameter variations. The discrete time small signal model can be written as

$$
\begin{bmatrix}
\delta x_{k+1} \\
\delta p_{k+1}
\end{bmatrix} =
\begin{bmatrix}
A_{1d} & A_{2d} \\
0 & I
\end{bmatrix}
\begin{bmatrix}
\delta x_k \\
\delta p_k
\end{bmatrix} +
\begin{bmatrix}
w_{1k} \\
w_{2k}
\end{bmatrix}
$$

\hspace{1cm}
(44)

$$
\delta y_k =
\begin{bmatrix}
C_1 & C_2
\end{bmatrix}
\begin{bmatrix}
\delta x_k \\
\delta p_k
\end{bmatrix} + e_k
$$

where the definitions of $A$ and $C$ are apparent from a comparison of the two preceding equations. Now we can use a Kalman filter to estimate $\delta x_k$ and $\delta p_k$. Actually, we are only interested in estimating $\delta p_k$ (the health parameter deviations), but the Kalman filter gives us the bonus of also estimating $\delta x_k$ (the excursions of the original turbofan state variables).

It is known that health parameters do not improve over time. That is, $\delta p(1)$, $\delta p(2)$, $\delta p(3)$, $\delta p(4)$, $\delta p(6)$, and $\delta p(8)$ are always less than or equal to zero and always decrease with time. Similarly, $\delta p(5)$ and $\delta p(7)$ are always greater than or equal to zero and always increase with time. In addition, it is known that the health parameters vary slowly with time. As an example, since $\delta \hat{p}(1)$ is the constrained estimate of $\delta p(1)$, we can enforce the following constraints on $\delta \hat{p}(1)$.

\begin{align*}
\delta \hat{p}(1) & \leq 0 \\
\delta \hat{p}_{k+1}(1) & \leq \delta \hat{p}_k(1) + \gamma_1^+ \\
\delta \hat{p}_{k+1}(1) & \geq \delta \hat{p}_k(1) - \gamma_1^-
\end{align*}

\hspace{1cm}
(46)

where $\gamma_1^+$ and $\gamma_1^-$ are nonnegative factors chosen by the user that allows the state estimate to vary only within prescribed limits. Typically we choose $\gamma_1^- > \gamma_1^+$ so that the state estimate can change more in the negative direction than in the positive
direction. This is in keeping with our *a priori* knowledge that this particular state variable never increases with time. Ideally we would have $\gamma_1^+ = 0$ since $\delta p(1)$ never increases. However, since the state variable estimate varies around the true value of the state variable, we choose $\gamma_1^+ > 0$. This allows some time-varying increase in the state variable estimate to compensate for a state variable estimate that is smaller than the true state variable value.

These constraints are linear and can therefore easily be incorporated into the form required in the constrained filtering problem statement (14). Note that this does not take into account the possibility of abrupt changes in health parameters due to discrete damage events. That possibility must be addressed by some other means (e.g., residual checking [3]) in conjunction with the methods presented in this paper.

## 6 Simulation Results

We simulated the methods discussed in this paper using MATLAB. We simulated a steady state 3 second burst of engine data measured at 10 Hz during each flight. Each of these routine services was performed at the single operating point shown in Tables 1–4. The signal-to-noise ratios were determined on the basis of NASA experience and previously published data [23] and are shown in Table 4. We used a one-sigma process noise in the Kalman filter equal to 1% of the nominal state values to allow the filter to be responsive to changes in the state variables. We set the one sigma process noise for each component of the health parameter portion of the state derivative equation to 0.01% of the nominal parameter value. This was obtained by tuning. It was small enough to give reasonably smooth estimates, and large enough to allow the filter to track slowly time-varying parameters. For the filter with hard constraints, we chose the $\gamma$ variables in (46) such that the maximum allowable rate of change in $\delta p$ was a linear 9% per 500 flights in the direction of expected change, and 3% per 500 flights in the opposite direction. The true health parameter values never change in a direction opposite to the expected change. However, we allow the state estimate to change in the opposite direction to allow the Kalman filter
to compensate for the fact that the state estimate might be either too large or too small. We set the weighting matrix $W$ in (23) and (31) equal to $\Sigma^{-1}$ in accordance with Theorem 3. We found by experimenting that setting the weighting matrix $V_k$ in (31) equal to $120W$ resulted in good performance for the Kalman filter with soft constraints.

The first test case we simulated was a linear degradation of the first health parameter (fan airflow) over 500 flights, while the other seven health parameters remained constant. Figure 2 shows the Kalman filters’ performances in this case. We ran eight simulations like this. In each simulation, one of the eight health parameters degraded linearly by a factor of 3% during the course of the simulation, while the other seven health parameters remained constant. The 3% degradation over 500 flights is in line with turbofan performance data collected by NASA and reported in the literature [25]. Each of the eight cases exhibit performance similar to Figure 2. Table 5 shows the performance of the filters averaged over all eight simulations. All of the filters estimate the health parameters to within less than 0.2% of their nominal values. It can be seen that (on average) the filter with soft constraints offers an 11% improvement over the unconstrained filter, and the filter with hard constraints offers a 22% improvement over the unconstrained filter. These numbers should not be interpreted as having any statistical significance (due to our limited sample size of eight cases) but they do show the improvement that is possible with constrained Kalman filters. Table 5 also shows that a couple of health parameters (fan airflow and LPT airflow) were actually estimated better with the unconstrained filter than with the constrained filter. We therefore see that the constrained filter does not guarantee better estimation in every individual sample run, but it does guarantee better performance statistically.

The next scenario we considered was the case where all eight health parameters degrade at the same time. We simulated a degradation over 500 flights of $-1\%$ for fan airflow, $-2\%$ for fan efficiency, $-3\%$ for compressor airflow, $-2\%$ for compressor efficiency, $+3\%$ for high pressure turbine airflow, $-2\%$ for high pressure turbine enthalpy change, $+2\%$ for low pressure turbine airflow, and $-1\%$ for low pressure turbine enthalpy change. This is summarized in Table 6. Figure 3 shows
the performance of the Kalman filters in this case. Table 7 shows the performance of the filters averaged over 16 simulations like this (each simulation being subject to a different random noise history). It can be seen that (on average) the filter with soft constraints offers a 9% improvement over the unconstrained filter, and the filter with hard constraints offers a 38% improvement over the unconstrained filter. As mentioned above, these numbers should not be interpreted as having any statistical significance (due to our limited sample size of 16 cases) but they do show the improvement that is possible with constrained Kalman filters.

The improved performance of the constrained filters comes with a price, and that price is computational effort. The filter with soft constraints requires only slightly (14%) more computational effort than the unconstrained filter, but the filter with hard constraints requires about four times the computational effort of the unconstrained filter. This is because of the additional quadratic programming problem that is required for hard constraints. However, computational effort is not a critical issue for the particular application of turbofan health estimation since the filtering is performed on ground-based computers after each flight.

7 Conclusion and Discussion

We have presented two methods for incorporating linear state inequality constraints in a Kalman filter. The first method incorporated hard constraints into the Kalman filter to maintain the state variable estimates within a user-defined envelope. The second method incorporated soft constraints into the Kalman filter to ensure that the state variable estimates vary slowly with time. The simulation results demonstrate the effectiveness of these methods, particularly for turbofan engine health estimation.

If the system whose state variables are being estimated has known state variable constraints, then those constraints can be incorporated into the Kalman filter as shown in this paper. However, in practice, the constraints enforced in the filter might be more relaxed than the true constraints. This allows the filter to correct state variable estimates in a direction that the true state variables might never
change. This is a departure from strict adherence to theory, but in practice this improves the performance of the filter. This is an implementation issue that is conceptually similar to tuning a standard Kalman filter.

It was seen in Theorem 2 that the filter with hard constraints has a smaller estimation error covariance than the unconstrained Kalman filter. At first this seems counterintuitive, since the standard Kalman filter is by definition the minimum variance filter. However, we have changed the problem by introducing state variable constraints. Therefore, the standard Kalman filter is not the minimum variance filter for the turbofan engine health estimation problem, and we can do better with the constrained Kalman filter.

We saw that the filter with hard constraints required a much larger computational effort than the standard Kalman filter. This is due to the addition of the quadratic programming problem that must be solved in the constrained Kalman filter. The engineer must therefore perform a tradeoff between computational effort and estimation accuracy. For real time applications the improved estimation accuracy may not be worth the increase in computational effort.

It was seen in Figures 2 and 3 that although the constrained filters improve the estimation accuracy, the general trend of the state variable estimates does not change with the introduction of state constraints. This is because the constrained filters are based on the unconstrained Kalman filter. The constrained filter estimates therefore have the same shape as the unconstrained estimates until the constraints are violated, at which point the state variable estimates are projected onto the edge of the constraint boundary. The constrained filters presented in this paper are not qualitatively different than the standard Kalman filter; they are rather a quantitative improvement in the standard Kalman filter.

Note that the Kalman filter works well only if the assumed system model matches reality fairly closely. The method presented in this paper, by itself, will not work well if there are large sensor biases or hard faults due to severe component failures. A mission-critical implementation of a Kalman filter should always include some sort of residual check to verify the validity of the Kalman filter results, particularly for the application of turbofan engine health estimation considered in this paper [3, 26].
Although we have considered only linear state constraints, it is not conceptually difficult to extend this paper to nonlinear constraints. If the state constraints are nonlinear they can be linearized as discussed in [2].

Further work along the lines of this research could focus on combining our work with [27] in order to guarantee convergence in the presence of nonlinear constraints. Other efforts could explore the incorporation of state constraints for optimal smoothing, or the use of state constraints in $H_\infty$ filtering [28]. Further work could also focus on integrating the nonlinear simulation logic in DIGTEM [8, 22] with the Kalman filter to obtain more complete results. This would also allow us to more easily test the Kalman filter at various operating points without translating data from DIGTEM to MATLAB.

References


Figure 1: Schematic representation of turbofan engine
Figure 2: Kalman filter estimates of health parameters. The true health parameter changes were a −3% change in the first parameter, and zero change in the other seven parameters. The true health parameter changes are shown as heavy lines, and the filter estimates are shown as lighter lines.
Figure 3: Kalman filter estimates of health parameters. The true health parameter changes were various values in between $-3\%$ and $+3\%$. The true health parameter changes are shown as heavy lines, and the filter estimates are shown as lighter lines.
<table>
<thead>
<tr>
<th>State</th>
<th>Nominal Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Pressure Turbine Rotor Speed</td>
<td>6140 RPM</td>
</tr>
<tr>
<td>High Pressure Turbine Rotor Speed</td>
<td>9395 RPM</td>
</tr>
<tr>
<td>Compressor Mass Flow</td>
<td>0.457 kg/s</td>
</tr>
<tr>
<td>Combustor Inlet Temperature</td>
<td>965 K</td>
</tr>
<tr>
<td>Combustor Mass Flow</td>
<td>0.264 kg/s</td>
</tr>
<tr>
<td>High Pressure Turbine Inlet Temperature</td>
<td>1593 K</td>
</tr>
<tr>
<td>High Pressure Turbine Mass Flow</td>
<td>1.48 kg/s</td>
</tr>
<tr>
<td>Low Pressure Turbine Inlet Temperature</td>
<td>1129 K</td>
</tr>
<tr>
<td>Low Pressure Turbine Mass Flow</td>
<td>1.79 kg/s</td>
</tr>
<tr>
<td>Augmentor Inlet Temperature</td>
<td>790 K</td>
</tr>
<tr>
<td>Augmentor Mass Flow</td>
<td>1.46 kg/s</td>
</tr>
<tr>
<td>Nozzle Inlet Temperature</td>
<td>790 K</td>
</tr>
<tr>
<td>Duct Fluid Momentum</td>
<td>53.6 kg/s²</td>
</tr>
<tr>
<td>Augmentor Fluid Momentum</td>
<td>103 kg/s²</td>
</tr>
<tr>
<td>Duct Mass Flow</td>
<td>4.52 kg/s</td>
</tr>
<tr>
<td>Duct Temperature</td>
<td>571 K</td>
</tr>
</tbody>
</table>

Table 1: Turbofan states.

<table>
<thead>
<tr>
<th>Control</th>
<th>Nominal Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Combustor Fuel Flow</td>
<td>0.37 kg/s</td>
</tr>
<tr>
<td>Augmentor Fuel Flow</td>
<td>0 kg/s</td>
</tr>
<tr>
<td>Nozzle Throat Area</td>
<td>430 cm²</td>
</tr>
<tr>
<td>Nozzle Exit Area</td>
<td>492 cm²</td>
</tr>
<tr>
<td>Fan Vane Angle</td>
<td>−25 deg</td>
</tr>
<tr>
<td>Compressor Van Angle</td>
<td>−20 deg</td>
</tr>
</tbody>
</table>

Table 2: Turbofan controls.
<table>
<thead>
<tr>
<th>Health Parameter</th>
<th>Nominal Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fan Airflow</td>
<td>102 kg/s</td>
</tr>
<tr>
<td>Fan Efficiency</td>
<td>0.82</td>
</tr>
<tr>
<td>Compressor Airflow</td>
<td>48.7 kg/s</td>
</tr>
<tr>
<td>Compressor Efficiency</td>
<td>0.83</td>
</tr>
<tr>
<td>High Pressure Turbine Airflow</td>
<td>41.0 kg/s</td>
</tr>
<tr>
<td>High Pressure Turbine Enthalpy Change</td>
<td>101 J/kg</td>
</tr>
<tr>
<td>Low Pressure Turbine Airflow</td>
<td>48.3 kg/s</td>
</tr>
<tr>
<td>Low Pressure Turbine Enthalpy Change</td>
<td>27.1 J/kg</td>
</tr>
</tbody>
</table>

Table 3: Turbofan health parameters.

<table>
<thead>
<tr>
<th>Measurement</th>
<th>Nominal Value</th>
<th>SNR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low Pressure Turbine Rotor Speed</td>
<td>6140 RPM</td>
<td>150</td>
</tr>
<tr>
<td>High Pressure Turbine Rotor Speed</td>
<td>9395 RPM</td>
<td>150</td>
</tr>
<tr>
<td>Duct Pressure</td>
<td>19.0 N/cm²</td>
<td>200</td>
</tr>
<tr>
<td>Duct Temperature</td>
<td>571 K</td>
<td>100</td>
</tr>
<tr>
<td>Compressor Inlet Pressure</td>
<td>20.5 N/cm²</td>
<td>200</td>
</tr>
<tr>
<td>Compressor Inlet Temperature</td>
<td>577 K</td>
<td>100</td>
</tr>
<tr>
<td>Combustor Pressure</td>
<td>97.5 N/cm²</td>
<td>200</td>
</tr>
<tr>
<td>Combustor Inlet Temperature</td>
<td>965 K</td>
<td>100</td>
</tr>
<tr>
<td>Low Pressure Turbine Inlet Pressure</td>
<td>26.8 N/cm²</td>
<td>100</td>
</tr>
<tr>
<td>Low Pressure Turbine Inlet Temperature</td>
<td>1130 K</td>
<td>70</td>
</tr>
<tr>
<td>Augmentor Inlet Pressure</td>
<td>17.4 N/cm²</td>
<td>100</td>
</tr>
<tr>
<td>Augmentor Inlet Temperature</td>
<td>790 K</td>
<td>70</td>
</tr>
</tbody>
</table>

Table 4: Turbofan measurements.
### Table 5: Kalman filter estimation errors. HPT = High Pressure Turbine, and LPT = Low Pressure Turbine. The numbers shown are RMS estimation errors (percent) averaged over eight simulations where each simulation had one health parameter degradation while the other seven health parameters were unchanged.

<table>
<thead>
<tr>
<th>Health Parameter</th>
<th>Unconstrained Filter</th>
<th>Soft Constrained Filter</th>
<th>Hard Constrained Filter</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fan Airflow</td>
<td>0.123</td>
<td>0.105</td>
<td>0.139</td>
</tr>
<tr>
<td>Fan Efficiency</td>
<td>0.177</td>
<td>0.166</td>
<td>0.113</td>
</tr>
<tr>
<td>Compressor Airflow</td>
<td>0.145</td>
<td>0.132</td>
<td>0.113</td>
</tr>
<tr>
<td>Compressor Efficiency</td>
<td>0.102</td>
<td>0.086</td>
<td>0.059</td>
</tr>
<tr>
<td>HPT Airflow</td>
<td>0.116</td>
<td>0.100</td>
<td>0.101</td>
</tr>
<tr>
<td>HPT Enthalpy Change</td>
<td>0.093</td>
<td>0.081</td>
<td>0.055</td>
</tr>
<tr>
<td>LPT Airflow</td>
<td>0.104</td>
<td>0.090</td>
<td>0.109</td>
</tr>
<tr>
<td>LPT Enthalpy Change</td>
<td>0.181</td>
<td>0.168</td>
<td>0.118</td>
</tr>
<tr>
<td>Average</td>
<td>0.130</td>
<td>0.116</td>
<td>0.101</td>
</tr>
</tbody>
</table>

Table 6: Health parameter degradation amounts for test scenario.

<table>
<thead>
<tr>
<th>Health Parameter</th>
<th>True Degradation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fan Airflow</td>
<td>-1%</td>
</tr>
<tr>
<td>Fan Efficiency</td>
<td>-2%</td>
</tr>
<tr>
<td>Compressor Airflow</td>
<td>-3%</td>
</tr>
<tr>
<td>Compressor Efficiency</td>
<td>-2%</td>
</tr>
<tr>
<td>HPT Airflow</td>
<td>+3%</td>
</tr>
<tr>
<td>HPT Enthalpy Change</td>
<td>-2%</td>
</tr>
<tr>
<td>LPT Airflow</td>
<td>+2%</td>
</tr>
<tr>
<td>LPT Enthalpy Change</td>
<td>-1%</td>
</tr>
</tbody>
</table>

Table 6: Health parameter degradation amounts for test scenario.
<table>
<thead>
<tr>
<th>Health Parameter</th>
<th>Estimation Error (%)</th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Unconstrained Filter</td>
<td>Soft Constrained Filter</td>
<td>Hard Constrained Filter</td>
</tr>
<tr>
<td>Fan Airflow</td>
<td>0.129</td>
<td>0.113</td>
<td>0.089</td>
<td></td>
</tr>
<tr>
<td>Fan Efficiency</td>
<td>0.163</td>
<td>0.149</td>
<td>0.105</td>
<td></td>
</tr>
<tr>
<td>Compressor Airflow</td>
<td>0.152</td>
<td>0.146</td>
<td>0.103</td>
<td></td>
</tr>
<tr>
<td>Compressor Efficiency</td>
<td>0.101</td>
<td>0.087</td>
<td>0.052</td>
<td></td>
</tr>
<tr>
<td>HPT Airflow</td>
<td>0.119</td>
<td>0.114</td>
<td>0.076</td>
<td></td>
</tr>
<tr>
<td>HPT Enthalpy Change</td>
<td>0.092</td>
<td>0.078</td>
<td>0.050</td>
<td></td>
</tr>
<tr>
<td>LPT Airflow</td>
<td>0.104</td>
<td>0.091</td>
<td>0.057</td>
<td></td>
</tr>
<tr>
<td>LPT Enthalpy Change</td>
<td>0.168</td>
<td>0.155</td>
<td>0.111</td>
<td></td>
</tr>
<tr>
<td>Average</td>
<td>0.128</td>
<td>0.116</td>
<td>0.080</td>
<td></td>
</tr>
</tbody>
</table>

Table 7: Kalman filter estimation errors. HPT = High Pressure Turbine, and LPT = Low Pressure Turbine. The numbers shown are RMS estimation errors (percent) averaged over 16 simulations, where each simulation had a linear degradation of all eight health parameters.