# Kalman Filters.

Bradley Thompson, Christopher Stone

April 10, 2002

# Contents

# 1 Introduction

## 1.1 What are filters

Filters are a way of removing something from a system to get a clean system. For example, a filter could remove suspended particle matter from a liquid to give a clean liquid. Filters can also be used to remove noise from a system, a simple example being a pair of earplugs.

## 1.2 Why do we use filters

When we measure the state of a system, our measurements are corrupted with noise. Systems themselves are inherently noisy, only mathematical equations are perfect. We use filters to remove the noise from the measurements and compensate for the noise in the system to a gain an accurate representation of the system.

## 1.3 History of filters

The Discrete-data linear filtering problem is that you get noisy measurements of a system measured over discrete time intervals, where the system can be represented by a linear equation. The best way of solving this problem until Kalman filters was to use a Wiener filter. Wiener filters had a number of limitations which reduced their use in real life applications, they needed to store all the previous states and used all of those previous states to calculate the estimate for the next state.

In 1960 Rudolph E. Kalman published a paper which described a minimal, recursive solution to the discrete-data linear filtering problem. This has lead to the Kalman filter becoming a popular and much worked on filter. The filter that was first published by Rudolph Kalman dealt only with discrete-data of a linear function, but this has been extended to allow for non-linear systems.

## 1.4 Introduction to Kalman filters

The purpose of a Kalman filter is to estimate the state of a system from measurements which contain random errors.

To start with, a Kalman filter requires that the initial state of the system be described and that you give it a starting point in the system. Once it has these conditions it then takes the system and produces an prediction of the next state of the system. A measurement of the predicted state is then made and the filter updates its prediction with the measurement, producing an estimate for the current state. Another prediction, based on the current, estimated, state is then made. This process then continues in a recursive cycle. The Kalman filter does not require all of the previous state measurements and estimations to be kept. It only needs to know its estimate and the measurement for that state. The filter also takes into account how much you trust either the measurements or the predictions. If you are more confidant of one than the other, then the less favoured one will have less influence on the filter. For example you are 100% confident in the model then the filter will ignore the measurements.

### 1.5 Extended Kalman filters

The Kalman filter introduced by Rudolph Kalman dealt with systems that could be described by a linear stochastic difference equation. This meant that if the process and (or) the measurement relationship was not linear you could not measure it, but with the introduction of Extended Kalman filters you can. An Extended Kalman Filter is one that linearizes about the current mean and covariance. Using a Extended Kalman filter in something like a Taylor series, you can linearizes around the current estimation using partial derivatives of the process and measurement functions to get the non-linear estimation[1]. By modifying the original equations for the linear Kalman filter, we can estimate the non-linear process and (or) measurement relationship. One point of note is that distributions of the various random variables are no longer normal after they have been linearized. The Extended Kalman filter is only an ad hoc state estimator that only approximates the optimality of Bayes's rule by linearizion[1].

## 2 The Discrete Kalman Filter

### 2.1 Stochastic Equations

The Kalman Filter is used to estimate the state $x \in \mathbb{R}^n$ of a discrete-time controlled process that is governed by the following linear stochastic difference equation

$$x_t = Ax_{t-1} + Bu_t + w_{t-1},$$

with a measurement $z \in \mathbb{R}^m$

$$z_t = Hx_t + v_t.$$

This means that the state, $x$, of the system at any point in time can be expressed as a vector containing $n$ elements of real ($\mathbb{R}$) numbers. A stochastic equation is one involving random variables. In the above equations,

- $x_t$ is the state of the system at time $t$, where $x \in \mathbb{R}^n$.
- A is a $n \times n$ matrix relating the previous state to the current state in absence of either a driving function or any process noise i.e. how does the current state influence the next state? Being a $n \times n$ matrix preserves the dimensionality of the state vector.
- $u_t$ is an optional control input, i.e. the driving function. $u \in \mathbb{R}^l$, is a vector of $l$ real numbers representing the aspects of the system that one is controlling. $l$ need not be equal to $n$.
- B is a $n \times l$ matrix which relates the control input to the current state. Being a $n \times l$ matrix further preserves the dimensionality of the state vector $x$.
- w, v : Random variables representing process and measurement noise. They are assumed to be independent of each other, white, with normal probability distributions.

  $$p(W) \sim N(O, Q)$$

  $$p(V) \sim N(O, R)$$

  That is, the probability distributions for w and v ( $p(W)$ and $p(V)$ ) are normal ($N$) with a mean of zero (0) and a covariance of $Q$ and $R$ respectively.

- Q is the process noise covariance.
- R is the measurement noise covariance.
- $z_t$ is the measurement of the state at time $t$. $z \in \mathbb{R}^m$ defines a state measurement of $m$ real numbers, where $m$ need not be equal to $n$.
- H is a $m \times n$ matrix relating the state to the measurement $z_t$. This preserves the dimensionality of the state measurement $(z)$ and defines which parts of the state are actually being measured. For example, in a system where the state $x_t$ comprises of a position and a velocity, we might only measure the position. Thus $x \in \mathbb{R}^2$ and $z \in \mathbb{R}^1$ and H will be a $1 \times 2$ matrix, say $|1\ 0|$. This would mean that the measurement is only related to the position part of the state vector.

### 2.1.1 IMPORTANT

The discrete Kalman filter applies to discrete-time controlled processes that can be expressed as a linear equation. In order to apply the Kalman filter, one must understand the system one is trying to apply it to. Knowledge of the system in question is required in order to construct a suitable state representation, as well as finding values for the relation matrices.

## 2.2 Kalman Filter Equations

The actual equations that are used in the recursive Kalman Filter are as follows:

$$\hat{x}_t^- = A\hat{x}_{t-1} + Bu_t \tag{1}$$

$$P_t^- = AP_{t-1}A^T + Q \tag{2}$$

$$K_t = P_t^- H^T (HP_t^- H^T + R)^{-1} \tag{3}$$

$$\hat{x}_t = \hat{x}_t^- + K_t(z_t - H\hat{x}_t^-) \tag{4}$$

$$P_t = (I - K_t H)P_t^- \tag{5}$$

- $\hat{x}_t^-$ is the prediction of the next state at time $t$. Is a $n \times 1$ matrix.
- A is the $n \times n$ state relation matrix.
- $\hat{x}_{t-1}$ is the estimate of the previous state, at time $t-1$.
- $u_t$ is the optional control input to the system the state. Is a $l \times 1$ matrix.
- B is the $n \times l$ control input – state relation matrix.
- $P_t^-$ is the amount of uncertainty in the predicted state. Is a $n \times n$ matrix.
- $P_{t-1}$ is the amount of uncertainty in the previous, estimated, state. Is a $n \times n$ matrix.
- $A^T$ is the transpose of matrix A.
- Q is the process noise covariance matrix.
- $K_t$ is the Kalman gain. Is a $n \times m$ matrix.

- H is the $m \times n$ measurement – state relation matrix.
- $H^T$ is the transpose of matrix H.
- R is the measurement noise covariance matrix.
- $\hat{x}_t$ is the state estimate at time $t$, updated from the state prediction for time $t$.
- $z_t$ is the measurement of the current state. Is a $m \times 1$ matrix.
- $P_t$ is the updated uncertainty in the estimate of the current state of the system.
- I is the identity matrix. Is a $n \times n$ matrix.

The Kalman Filter equations (1-5) are split into Time Update ("Prediction") and Measurement Update ("Correction") sections. Equations (1) and (2) project the state and error covariance ahead one step. Equations (3-5) take the predictions and update the state estimate and error covariance, comparing the predicted state with the measurement of the actual state. We will now take a closer look at the equations, and then see how the whole thing works.

We talk about state estimates ($\hat{x}_t$), predictions ($\hat{x}_t$), and measurements ($z_t$) because the true state of the system is never actually known due to the presence of noise. We make a prediction about the next state of the system, then measure that state directly (with associated noise) to "correct" the prediction, giving us an estimate for the current state of the system.

#### 2.2.1 Prediction

$$\hat{x}_t^- = A\hat{x}_{t-1} + Bu_t$$

$$P_t^- = AP_{t-1}A^T + Q$$

These two equations are responsible for taking the current state and projecting forward in time an estimate for the next state. The uncertainty in the prediction is also calculated and projected forward.

Equation (1) calculates the next state $\hat{x}_t^-$ from the previous state $\hat{x}_{t-1}$ and any inputs $u_t$. The matrices $A$ and $B$ determine how the previous state and inputs relate to the next state in that $A$ determines how each state is related to the one before e.g. is each state equal to the previous state or is each state two times the previous state?. $B$ determines which parts of the state vector are adjusted by the input, and by how much.

Equation (2) calculates a level of uncertainty for the state prediction. The uncertainty is calculated from the previous states updated uncertainty, with the addition of process noise in the form of the covariance matrix $Q$.

#### 2.2.2 Correction

$$K_t = P_t^- H^T (HP_t^- H^T + R)^{-1}$$

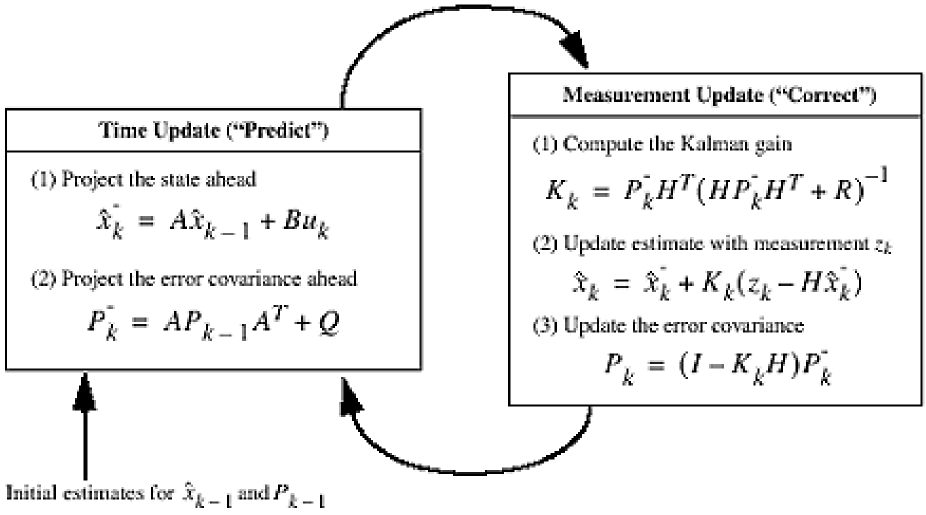


Figure 1: A complete picture of the Kalman Filter. From SIGGRAPH 2001.

$$\hat{x}_t = \hat{x}_t^- + K_t(z_t - H\hat{x}_t^-)$$
$$P_t = (I - K_t H)P_t^-$$

These three equations take the estimated state and uncertainty and "correct" them based on a measurement taken of the actual state.

Equation (3) calculates the Kalman gain for the state.

Equation (4) updates the state prediction using the calculated Kalman gain and the state measurements to form an estimation of the current state of the system.

Equation (5) uses the Kalman gain to update the uncertainty of the prediction. The new state estimation uncertainty will be better or worse depending on how good the original prediction was.

### 2.2.3 Prediction – Correction Cycle

To use the Kalman filter, one must first describe the system in terms of a linear stochastic difference equation. That is, the matrices $A$, $B$ and $H$, the input vector $u$, and the process and measurement noise covariances $Q$ and $R$ must be known and set up. Initial values for $\hat{x}_0$ and $P_0$ must also be chosen as input into the prediction equations. The initial values need not be correct, or even accurate in any way – they are just guesses.

Once the filter has started, the prediction equations will predict the next state $\hat{x}_1^-$ and uncertainty $P_1^-$. The process is then repeated, forming a Prediction – Correction cycle, the essence of the Kalman filter. Figure 1 shows the Kalman filter cycle in action.

# 3 Examples

## 3.1 A simple example

The following example has been slightly modified from the one presented at SIGGRAPH 2001.

### 3.1.1 The Process Model

In this simple example let us attempt to estimate a scalar random constant, a voltage for example. We assume that we have the ability to take measurements of the system, i.e. with a voltmeter, but that the measurements are corrupted by a 0.1 volt RMS white measurement noise. We note the following:

- We are estimating a constant so the system does not change from state to state, so $A = 1$. i.e. the next state is $1 \times$ current state.
- There is no control input so $u = 0$. i.e we are not changing the system in any way.
- Our noisy measurement is of the state directly so $H = 1$.

The stochastic linear difference equation for this process thus simplifies to

$$x_t = x_{t-1} + w_{t-1}$$

with a measurement $z \in \mathbb{R}^1$

$$z_t = x_t + v_t$$

### 3.1.2 The filter equations

The Time-Update and Measurement-Update equations also simplify to become:

$$\hat{x}_t^- = \hat{x}_{t-1}, \tag{6}$$

$$P_t^- = P_{t-1} + Q, \tag{7}$$

$$K_t = P_t^-(P_t^- + R)^{-1}, \tag{8}$$

$$\hat{x}_t = \hat{x}_t^- + K_t(z_t - \hat{x}_t^-), \tag{9}$$

$$P_t = (I - K_t)P_t^-. \tag{10}$$

We presume a very small process variance, since it is supposed to be a constant, so we let $Q = 0.0001$. (Q could be zero, but giving it a non-zero value gives the filter more room to move). We also assume that the true value of our constant has a standard normal probability distribution, and knowing that the true value is five, we "seed" the filter with the guess that the constant is four. i.e. let $\hat{x}_{t-1} = 4$. This is a guess, it could be anything we want, but choosing a value far removed from the true one will mean the filter will take longer to converge. We choose a close value for simplicities sake.

We must also choose an initial value for $P_{t-1}$, based on how correct we think our initial guess is. If we were absolutely certain that our initial guess was correct, we could make $P_{t-1} = 0$. Unfortunately, this would make our filter initially and always believe $\hat{x}_{t-1} = 4$, fine if the guess was correct, but not much use otherwise. The alternative choice is not critical though, so, to be simple, we set $P_{t-1} = 1$.

We noted above that our measurements of the process are corrupted by a 0.1 RMS white noise, so we set the measurement noise covariance to $R = 0.01$. i.e. the measurements have noise with a standard deviation of 0.1, and thus a covariance of 0.01.

We then run the Kalman filter program (Appendix A), with the parameters as shown in Table 1, on fifty values normally distributed about the value five (Appendix B). The resulting correction to the Kalman filter state estimation is shown in Fig. 2.

| *parameter* | *value* |
|---|---|
| A | 1 |
| $u_t$ | 0 |
| H | 1 |
| Q | 0.0001 |
| R | 0.01 |
| $x_0$ | 4 |
| $P_0$ | 1 |

Table 1: Summary of parameters for running the Kalman filter for Example 1

### 3.2 A harder example

The next example is of a simple mathematic equation.

#### 3.2.1 The Process Model

In this example let us attempt to estimate a the simple equation $y = x$. As the system is a known equation we can produce the results, but we will introduce a little noise to the system. The measurements of the system will be corrupted by a noise of up to $\pm 4$.

- The states are related directly to each other, so $A = 1$, but now we have a driving function.
- There is a control input of $+1$, so $u_t = 1$. i.e a state is the previous state + 1.
- Our noisy measurement is of the state directly so $H = 1$.

The stochastic linear difference equation for this process thus simplifies to:

$$x_t = x_{t-1} + u_t + w_{t-1}$$

with a measurement $z \in \mathbb{R}^1$

$$z_t = x_t + v_t$$

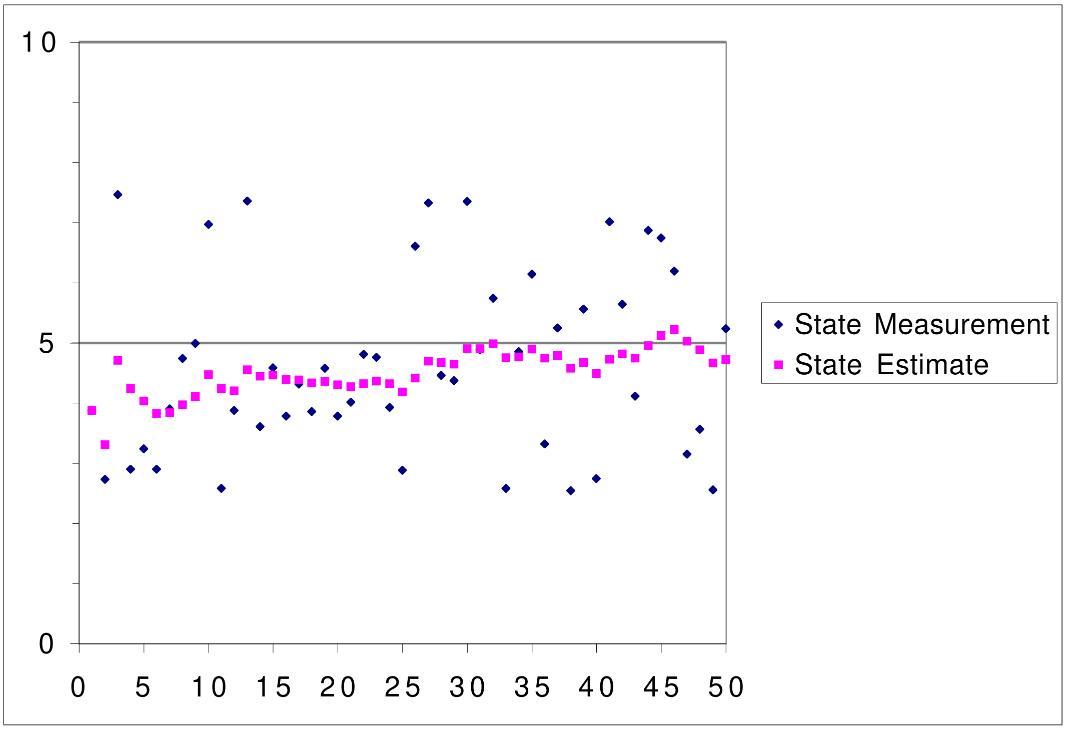


Figure 2: The graph shows the state estimate converging about the known value of the system, despite the noisy measurements

### 3.2.2 The filter equations

The Time-Update and Measurement-Update equations also simplify to become:

$$\hat{x}_t^- = \hat{x}_{t-1} + u_t, \tag{11}$$

$$P_t^- = P_{t-1} + Q, \tag{12}$$

$$K_t = P_t^-(P_t^- + R)^{-1}, \tag{13}$$

$$\hat{x}_t = \hat{x}_t^- + K_t(z_t - \hat{x}_t^-), \tag{14}$$

$$P_t = (I - K_t)P_t^-. \tag{15}$$

We presume a very small process variance, since it is increasing at a constant rate, so we let $Q = 0.0001$. (Q could be zero, but giving it a non-zero value gives the filter more room to move). We also assume that the true value of our state has a standard normal probability distribution, and the start point of the equation is 0. We "seed" the filters start point as before with the value twenty. i.e. let $\hat{x}_{t-1} = 20$. Again by choosing a value removed from the true one will mean the filter will take longer to converge.

We must also choose an initial value for $P_{t-1}$, based on how correct we think our initial guess is. We again set $P_{t-1} = 1$.

We set the measurement noise covariance to $R = 0.01$.

We then run the Kalman filter program (Appendix A), with the parameters as shown in Table 2, on fifty values normally distributed about the equation $y = x$ (Appendix D). The resulting correction to the Kalman filter state estimation is shown in Fig. 3.

| *parameter* | *value* |
|---|---|
| A | 1 |
| $u_t$ | 1 |
| H | 1 |
| Q | 0.0001 |
| R | 0.01 |
| $x_0$ | 20 |
| $P_0$ | 1 |

Table 2: Summary of parameters for running the Kalman filter for Example 2

## 3.3 An even harder example

In this example we will look at an accelerating object.

### 3.3.1 The Process Model

In this example let us attempt to estimate an object falling due to gravity. As the system is known we can produce the results, but we will introduce noise to the system.

- The current state is now related to the previous by the matrix:

$$A = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix}$$

- There is a control input so $u = -gravity$. i.e the system is the previous state $-$ gravity as the body id falling.

- There is relation as to how the control input effects the state. i.e

$$B = \begin{vmatrix} 0.5 \\ 1 \end{vmatrix}$$

- Our noisy measurement is of the state directly so

$$H = \begin{vmatrix} 1 \\ 0 \end{vmatrix}$$

The stochastic linear difference equation for this process thus simplifies to

$$x_t = Ax_{t-1} + Bu_t + w_{t-1}$$

with a measurement $z \in \mathbb{R}^1$

$$z_t = Hx_t + v_t$$

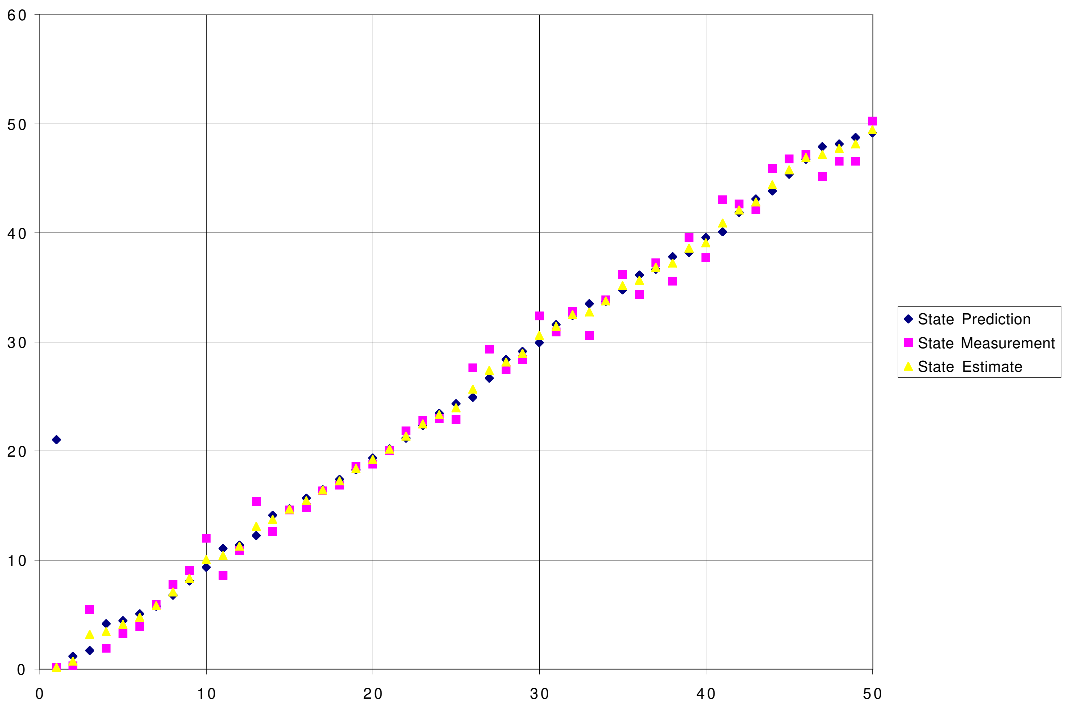


Figure 3: The graph shows the state prediction and state estimate variables mirroring the function. The initial prediction of 20 is essentially ignored

### 3.3.2 The filter equations

The Time-Update and Measurement-Update equations also simplify to become:

$$\hat{x}_t^- = A\hat{x}_{t-1} + Bu_t, \tag{16}$$

$$P_t^- = AP_{t-1}A^t + Q, \tag{17}$$

$$K_t = P_t^- H^t (HP_t^- H^t + R)^{-1}, \tag{18}$$

$$\hat{x}_t = \hat{x}_t^- + K_t(z_t - H\hat{x}_t^-), \tag{19}$$

$$P_t = (I - K_t H)P_t^-. \tag{20}$$

The state representation $x_t$ now has two variable the velocity and the height of the object. We presume a very small process variance so Q = 0.0001. The measurement noise covariance R = 0.01. Start off $x_0 = 495$. $P_0 = 1$

| *parameter* | *value* |
|---|---|
| A | $\begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix}$ |
| $u_t$ | -9.8 |
| H | $\left\|\begin{matrix} 1 \\ 0 \end{matrix}\right\|$ |
| Q | 0.0001 |
| R | 0.01 |
| $x_0$ | 495 |
| $P_0$ | 1 |

Table 3: Summary of parameters for running the Kalman filter for Example 3

# 4 Closing

## 4.1 Applications

The main application for Kalman filters is in measuring a system and in which you wish to minimize the amount of error due to noise. For example you might want to measure the voltage across a resistor. By adding a Kalman filter to the process you can increase your measurements accuracy by removing the noise in the measurements caused by random fluxuations in the voltage.

## 4.2 Applications in computer Vision

One of the tasks of computer vision is measuring something in the real world. As stated before, every measurement has random noise in it. The Kalman filter would be employed to remove the noise from the system to give cleaner measurements of that real world. The Kalman filter has also been implemented and developed in tracking for interactive graphics and motion tracking[1]. It can be used anywhere there is measurements. A good example of Kalman filters at work is in the University of Otago's Graphics department's Watching Windows. In this it is used to estimate between the frames where the eyes and hands of the subjects are.

### 4.3 Conclusion

As you can see from the examples a Kalman filter is able to estimate a system to a high level of accuracy given a system that has noise in it. From this we can also say that a Kalman filter also preforms well in a linear system and that it can be extend to operate on a non-linear system. The Kalman filter has only five equations and can be implemented easily as long as the system that it is going to be estimated is well known by the user.

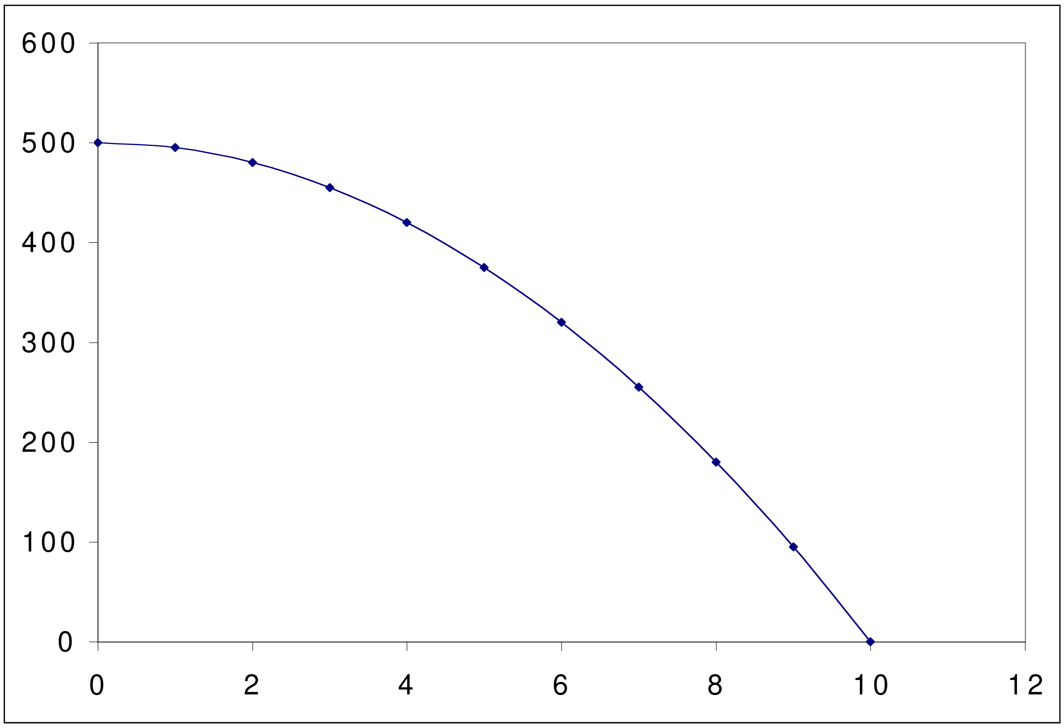


Figure 4: A Bodies acceleration due to gravity.

# 5 Glossary

**Discrete Data :** A finite or countable set of values, not continuous.

**Covariance :** A statistical measure of the variance of two random variables that are observed or measured in the same mean time period. This measure is equal to the product of the deviations of corresponding values of the two variables from their respective means.

**Stochastic :** Involving or containing a random variable or variables

**State Relation Matrix :** Relates one state to the next state. For example $x_t = 5 \times x_{t-1}$, and the State Relation matrix is 5. With a matrix

$$\begin{vmatrix} d_t \\ v_t \end{vmatrix} = \begin{vmatrix} 1 & 1 \\ 0 & 1 \end{vmatrix} \begin{vmatrix} d_{t-1} \\ v_{t-1} \end{vmatrix}$$

**Control input :** This is a control on the system, for example gravity.

## References


[1] Greg Welch, Gary Bishop: *An Introduction to the Kalman Filter* , Course proceedings from SIGGRAPH 2001.

[2] R. E. Kalman: *A New Approach to Linear Filtering and prediction Problems*, Transactions of the ASME–Journal of Basic Engineering, Volume 82, Series D, Pages 35-45, 1960